

Human Factors Study of CC-130 Operations



DEFC QUALITY INSPECTED 4

19970402 030

Defence and Civil
INSTITUTE OF ENVIRONMENTAL MEDICINE
and
Air Transport Group

Canada

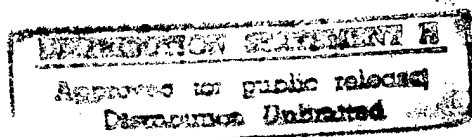
R.D. Banks
K.C. Hendy
W.D. Fraser
M.M. Thompson
D. Jamieson
H. Wright
T. Gee
I.M. Mack
J. Latulippe
B. Davis
M. Cole

December 1996

DCIEM No. 96-R-66

**DEFENCE AND CIVIL INSTITUTE OF
ENVIRONMENTAL MEDICINE
AND
AIR TRANSPORT GROUP
HUMAN FACTORS STUDY OF
CC-130 OPERATIONS**

R.D. Banks
K.C. Hendy
W.D. Fraser
M.M. Thompson
D.W. Jamieson
H. Wright
T. Gee
I.M. Mack
J. Latulippe
B. Davis
M. Cole



Defence and Civil Institute of Environmental Medicine
1133 Sheppard Avenue West, P.O. Box 2000
North York, Ontario
Canada M3M 3B9

© HER MAJESTY THE QUEEN IN RIGHT OF CANADA (1996)
as represented by the Minister of National Defence

© SA MAJESTE LA REINE EN DROIT DU CANADA (1996)
Défense Nationale Canada

DEPARTMENT OF NATIONAL DEFENCE – CANADA

ABSTRACT

The Defence and Civil Institute of Environmental Medicine (DCIEM) and Air Transport Group (ATG) were tasked to conduct a joint study of human factors concerning the CC-130 Hercules aircraft. The aim of the study was to establish human factors issues relevant to air accidents, and to recommend preventative measures. The study was organized around two working groups: the Crew Behaviour Assessment Group (CBAG) and the Flight Performance Assessment Group (FPAG). The CBAG developed a method of measuring the ability of the crew to coordinate their activities efficiently and manage their workload. The FPAG developed a method of measuring the accuracy and consistency of simulator flight along an aircraft flight path. Data to support the development of both methods were obtained from a simulator study of 23 ATG crews. The results defined the characteristics of high proficiency Aircraft Commanders (ACs) and those of less proficient ACs. Less proficient ACs seemed to focus primarily upon systems-related, procedural cross-checking and rechecking of information, and had more open-loop communication which supports the contention that these individuals were becoming task overloaded. The results suggest that a proportion of ATG crews are adversely overloaded by the occurrence of unexpected flight events and certain system failures. Since behaviour can be influenced by training, this study recommends a review and modification of the current CC-130 training program, including Aircrew Coordination Training (ACT).

EXECUTIVE SUMMARY

The Defence and Civil Institute of Environmental Medicine (DCIEM) and Air Transport Group (ATG) were tasked to conduct a joint study of human factors specific to safe flight operation of the CC-130 Hercules aircraft. The study was prompted by an apparently high accident rate in the CC-130 community. The aim of the study was to establish human factors issues relevant to air accidents, and to recommend preventative measures. The study was organized around two working groups: the Crew Behaviour Assessment Group (CBAG) and the Flight Performance Assessment Group (FPAG). Two methods were developed and applied independently by the groups. The CBAG developed a method of measuring the ability of the crew to coordinate their activities efficiently and manage their workload. The FPAG developed a method of measuring the accuracy and consistency of simulator flight along an aircraft flight path. Data to support the development of both methods were obtained from a simulator study of 23 ATG crews.

Results of the Crew Behaviour Study. The CBAG developed a measurement battery that has proven reliable, capable of yielding scientifically defensible results based on theory, and is applicable to a wide range of operational issues. In the process of developing this battery, the behavioural characteristics of highly and less proficient Aircraft Commanders (ACs) and crews were determined. The results indicated that **highly proficient** ACs were characterized by:

- a strong knowledge of systems and procedures,
- a greater likelihood to demonstrate a superior range and depth of thought concerning important aspects of the flight,
- the ability to address aspects of the flight that are more discretionary, such as weather and the mission, and
- greater resource/workload management skills both at their own individual level, and at the level of the team;

while **less proficient** ACs were characterized by:

- less knowledge of the CC-130 systems,
- less range and depth of thought (i.e., less preplanning), and
- evidence of work or information overload.

Less proficient ACs seemed to focus primarily upon systems-related, procedural cross-checking and rechecking of information. The co-pilots and the flight engineers of the less proficient ACs attempted **compensatory behaviours**, but this did not fully compensate for the deficits of these ACs. As well, there was a higher incidence of open-loop communication (communications left unanswered, unchallenged, or not acknowledged) for these individuals which supports the contention that these individuals were becoming task overloaded.

Results of the Flight Performance Study. Problems with data retrieval precluded definitive findings by the FPAG, but there was a suggestion that crews who managed resources effectively flew the aircraft most accurately and consistently. The results do indicate that the application of the developed tool, with large data sets, might provide a base for development of an in-flight safety monitoring capability.

Conclusions. While it is not possible to directly link the CC-130 accidents to the deficiencies identified in these studies, or indeed to any single common factor such as fatigue or experience, the results suggest that a proportion of ATG crews are adversely overloaded by the occurrence of unexpected flight events and certain system failures. Since behaviour can be influenced by training, a goal of the training program should be elimination of the ineffective behaviours seen in the less proficient crews, and the realization or reinforcement of those behaviours seen in the highly proficient crews. With this goal in mind, this study recommends a review and modification of the current CC-130 training program, including Aircrew Coordination Training (ACT).

TABLE OF CONTENTS

ABSTRACT	i
EXECUTIVE SUMMARY	iii
TABLE OF CONTENTS	v
LIST OF TABLES AND FIGURES.....	vii
TABLE OF ABBREVIATIONS	viii
INTRODUCTION.....	1
Background.....	1
Overall Study Aim.....	1
General Approach.....	1
Summary of Experimental Phase.....	2
Scope of this Report.....	2
ASSESSMENT OF CREW BEHAVIOUR.....	5
Introduction	5
Approach.....	5
Theoretical Framework	6
Method	7
Analysis.....	10
Results	10
Discussion.....	13
Implications of Findings to ATG.....	14
RECOMMENDATIONS.....	17
REFERENCES.....	19
ANNEX A. FLIGHT PERFORMANCE ASSESSMENT	21
ANNEX B. THEORETICAL BASIS FOR ASSESSING CREW BEHAVIOUR... 25	
Appendix 1. Communication and Crew Behaviour Coding.....	29
ANNEX C. ATG/DCIEM HUMAN FACTORS STUDY OF CC 130	
OPERATIONSFEASIBILITY AND PROGRESS 20 SEPT 1994	37
Appendix 1. Organization of the Study	45
Appendix 2. Work Plan	46
Appendix 3. Financial and Personnel Requirements	47

TABLE OF CONTENTS (continued)

ANNEX D. 1630-2 (HPSD) 15 MARCH 1995 ATG/DCIEM HUMAN FACTORS STUDY OF CC 130 OPERATIONS: PROGRESS REPORT	49
Appendix 1. Organization of the Study	57
Appendix 2. Work Plan	58
ANNEX E. REPORT OF THE BOX TOP STUDY.....	59
Appendix 1. Boxtop 1/95 Crew Callout Schedule	75
Appendix 2. Activity Log.....	76
Appendix 3. Sample Actigraph Plot.....	79

LIST OF TABLES AND FIGURES

Table 1. List of system malfunctions, flightpath and weather updates in the CC-130 flight simulator scenario.....	9
Table 2. Pattern of results of mental model domains and functions among highly and less proficient aircraft commanders.....	11
Table 3. Pattern of results for additional coding categories.....	12
Figure A1. Example of the flight scoring method.....	22
Table B1. Mental model domains and functions.....	27

TABLE OF ABBREVIATIONS

AC	Aircraft Commander
ACT	Aircrew Coordination Training
AIRCOM	Air Command
ATC	Air Traffic Control
ATG	Air Transport Group
CBAG	Crew Behaviour Assessment Group
CF	Canadian Forces
CFB	Canadian Forces Base
CP	Co-pilot
CRAD	Chief Research and Development
CRM	Crew Resource Management
DCIEM	Defence and Civil Institute of Environmental Medicine
FE	Flight Engineer
FPAG	Flight Performance Assessment Group
ICP	Instrument Check Pilot
IFR	Instrument Flight Rules
ILS	Instrument Landing System
Medevac	Medical Evacuation
NASA	National Aeronautics and Space Administration
SAR	Search and Rescue
SME	Subject Matter Expert

INTRODUCTION

Background

1. The Defence and Civil Institute of Environmental Medicine (DCIEM) and Air Transport Group (ATG) were tasked in January 1994 by Air Command (AIRCOM) to conduct a joint study of human factors specific to the Canadian Forces (CF) CC-130 Hercules operation (1). The study was prompted by a high accident rate in the CF CC-130 community.
2. The initial efforts consisted of a review of literature, a review of accident boards, and a survey of squadron attitudes regarding the priority of operations (2). Based on this initial review and on a series of working meetings, a study objective was defined, the literature review expanded, and a method developed. The general approach, as outlined in Reference 2, was approved in September 1994 and early progress reported in March 1995 (3). Two collateral studies investigated aircrew fatigue and the use of eye movement data as an aid to training. A study of fatigue and flight scheduling conducted during Exercise Box Top was reported previously (4).
3. The study was designed around two working groups: 1) The Crew Performance Assessment Group, renamed Crew Behaviour Assessment Group (CBAG); and 2) the Flight Performance Assessment Group (FPAG). The activities of these two working groups were supervised by a study group that included the chair, co-chair, heads of each working group, contractors, ATG, AIRCOM and Chief of Research and Development (CRAD) representatives. Data were collected during experiments that occurred between December 1995 and April 1996. Analysis of these data has constituted the major activity since April 1996. A briefing was given to the Commander ATG in July 1996 detailing the results of the study to date and observations made during the course of experiments conducted in the first half of the year. This report documents the material that supported this briefing and recommends the way ahead.

Overall Study Aim

4. The aim of the study was to establish human factors issues that may have contributed to an apparent increase in the incidence of fatal air accidents in CC-130 operations, and to recommend measures to prevent future accidents.

General Approach

5. In order to obtain scientifically defensible results, it was necessary to develop methods to collect and analyze data that would provide a basis for the study findings. As described previously by this study group (2), crew behaviour during flight operations is a process that can be observed. Also, movement of the aircraft through the atmosphere is a process that can be recorded through

flight data systems. A method was developed to measure 'outputs' from each of these different processes. The underlying assumption was that both processes relate to flight safety and operational effectiveness. Metrics were developed to score crew behaviour and the crew's ability to maintain target aircraft flight parameters.

6. Methods were developed independently by the two study groups so that an unbiased comparison of the results of the two methods might be made. The CBAG developed a metric of a crew's ability to coordinate their activities efficiently and manage their workload. The FPAG developed a metric that determined the accuracy and consistency of simulator flight along an aircraft flight path. As initially conceived, a combined method would consolidate these two tools. The consolidated method would then be used to examine select human factors issues.

7. The approach to measuring crew behaviour evolved into one that developed measures for assessing safe flight performance by identifying distinctive behaviours in highly proficient crews. These measured behaviours were then compared against those of less proficient crews. Since behaviours can be altered through training, the results of the crew behavioural study have implications for the training system and provide the major basis of the recommendations of this study.

Summary of Experimental Phase

8. The experimental phase of this study took place in the CC-130 Flight Simulator located at CFB Trenton. Flight parameter data and video and audio records were taken and transported to DCIEM for analysis. A total of 23 crews consisting of an Aircraft Commander, Co-pilot and Flight Engineer participated. Crew behaviour measurements were based upon the analyses of the video and audio recordings, and on direct observation of each simulator flight.

Scope of this Report

9. The results of the CBAG have significant implications for ATG training, standards and CRM issues. The results of the FPAG, however, were limited by CC-130 simulator hardware problems which led to significant loss of data. While the results suggested that crews who managed their resources effectively flew most accurately, the findings are of limited use. For that reason, the main body of this report is concerned with CBAG findings. The FPAG report and recommendations are found at Annex A.

10. While the body of this report deals primarily with the methods and recommendations of the CBAG study, two collateral activities were pursued in addition to the main work of the CBAG and FPAG. These were:

- specific studies into fatigue and scheduling in ATG operations, and
- the use of eye movement technology as a potential aid to training instrument scan patterns and locus of attention.

These activities were either exploratory in nature (eye movement technology), and therefore are not reported in detail here, or have been reported separately (e.g., the report on the Box Top study (4) — see also Annex E). It is assumed that the importance of continuing studies into the issue of fatigue needs no further justification, and a recommendation to this effect will be made. The eye movement instrumentation has been demonstrated to 426 Squadron and sufficient interest was shown to recommend that further investigations be carried out to see if this technology could be integrated usefully into the training system.

ASSESSMENT OF CREW BEHAVIOUR

Introduction

11. The task of the CBAG was to develop a generic measurement battery for assessing behavioural aspects of safe crew performance. The approach incorporated the central components of two theoretical models of human information processing with the literature on team performance and crew resource management. Data to support the development of the behavioural measures were obtained from a study of 23 ATG crews in the motion-based simulator at CFB Trenton. As demonstrated here, the measurement battery is reliable, capable of yielding scientifically valid results based on theory, and is applicable to a wide range of issues of operational concern to ATG (the effects of fatigue or decreasing experience levels upon decision making, crew coordination etc.).

Approach

12. Once the conceptual aspects of the behavioural measures were decided, the measurement battery was refined and partially validated in a simulator study. This study involved videotaping crews flying a simulated medical evacuation mission. This mission was incorporated into their continuation training in the CC-130 flight simulator. The flight simulator was ideal as it allowed for important aspects of naturalistic decision-making environments, for instance: 1) dynamic, changing and unfolding requirements; 2) shifting, ill-defined or competing goals; and 3) action/feedback loops. Moreover, the simulator afforded greater scientific rigor and control than would have been possible in the actual aircraft, and greater realism than would have been possible in a laboratory experiment.

13. This study adopted a naturalistic decision-making research strategy by attempting to identify those characteristics which are the hallmarks of proficiency among ATG aircrew. In order to identify these characteristics, check pilots rated each of 23 crews on their proficiency, finally choosing 6 highly, and 6 less proficient crews for further analysis. It was expected that highly proficient versus less proficient crews would exhibit important and consistent differences in the categorization of communication patterns related to decision making, workload, and resource management, as they dealt with the challenges caused by various systems malfunctions, destination airport radar failure and changing weather conditions.

14. This type of methodology served two important purposes: 1) it provided an efficient way to test and refine our measurement battery; and 2) it provided important diagnostic information about the behaviours that particularly distinguish highly from less proficient aircrew.

Theoretical Framework

15. Contractors and DCIEM staff completed an extensive review of the literature, including Crew Resource Management (CRM) and the theoretical literature on crew behaviour and decision making. DCIEM had previously completed several years of theoretical study that was directly relevant to the design of this study. The theoretical basis for work presented in this report is summarized at Annex B.

16. As noted at Annex B, a key component of efficient decision making is the quality of the **mental model** that the individual holds. A mental model is defined as the knowledge necessary to perform a task and may encompass past, present, and future flight parameters, goals, and considerations. Well-developed mental models lead to more efficient information processing, decreased time pressure and workload, and better performance. A mental model **domain** refers to the discernible and distinguishable content areas of an individual's thoughts concerning a flight. In the experimental scenario the following domains were central:

- aircraft systems,
- procedures and checklists,
- geography or air picture,
- the mission, and
- the changing weather.

The **function** of a mental model refers to the process involved in performing a task and is also manifested in the verbal communications among the crew. Functions progress in complexity from a simple awareness of the state of the world to the development and implementation of plans to cope with that state. The measurement of these processes amounts to a functional analysis of cockpit communications.

17. Mental model domains and functions can, of course, be considered together. Moreover, one can think of domain and function as reflecting the range and the depth of the mental model respectively. For instance, the number of mental model domains considered during a flight is indicative of the range of thought demonstrated by an individual. Similarly, simple awareness statements, such as one indicating awareness of a system malfunction, would be classified as requiring less depth of thought than a statement noting the implications of a system malfunction, or a statement indicating preplanning in light of the implications of the failure. Table B1 of Annex B illustrates the relationship of content domain (range) and function (depth). Both the range and depth of the mental model provide candidate categories for a measurement battery.

18. Mental models become even more complicated when a task is to be

completed by a team of individuals. Researchers in the area suggest that it is overlap in shared mental models that is chiefly responsible for the consequent effectiveness or the lack of effectiveness of a team. A major function of both verbal and non-verbal communication is to build common mental models amongst the crew.

19. Prior research and our own preliminary observations led us to include the following additional behaviour categories into the measurement battery. The first is a category termed **systems knowledge**. Although fairly self-explanatory, this category would be used only when aircrew demonstrate that they knew exactly and immediately how to deal with a system malfunction, prior to consulting any checklists or reference manuals. Two further categories relate most directly to resource management skills and are termed **task prioritization** and **crew monitoring**. Although the former category requires no further explanation, the latter category refers to instances in which a crew member (most likely the Aircraft Commander or AC) actively and closely monitors other crewmembers' work and stress levels or their progress on a specific demanding flight task. A final category, referred to as **open loop communication**, would be used in instances in which a crewmember failed to respond to another crewmember's statement or query. This category is an important one as it signals a lack of crew communication. It is also a relatively good proxy measure for that crewmember's level of workload at that point in time. In essence, the crewmember simply does not have the resources to respond to all the inputs and demands at that moment. (See Appendix 1 of Annex B for coding categories).

Method

20. **Subjects.** Participants were drawn from each of the ATG CC-130 Squadrons. A total of 23 crews, each consisting of an AC, Co-pilot (CP), and Flight Engineer (FE) undergoing normal CC-130 simulator continuation training participated in the study. This is a significant number as it represents approximately one third of CC-130 crews.

21. During the flight task, the simulator instructor played the role of ATC, loadmaster, and any additional staff as required. There was no navigator on the flight as the simulator does not provide for this crew position. The absence of the Navigator was built into the details of the scenario under which the flight was conducted. The experimenter flew all simulator sessions but did not interact with the crews during the flight itself.

22. **Procedure.** Each crew arrived at the simulator and completed preliminary preparations for a local 'trainer' flight in the simulator. A Navigator is not normally carried on such flights. Just prior to the beginning of the simulator session the crew was brought into the briefing room and asked to participate in the ATG/DCIEM study. None of the 23 crew members refused to participate in the study. After their agreement, the experimenter provided a

verbal introduction to the study. The crew was informed that the purpose of the study was to observe ATG crews and to develop a sensitive metric of crew performance. Participants were told that, with their permission, their simulator session would be videotaped for later review at DCIEM, were assured of the confidentiality of their videotapes, and were asked if they had any questions about the study in general.

23. Next, participating aircrew were told that the nature of their mission had been altered. Instead of the local trainer they had expected to fly, they would fly a medical evacuation (medevac) mission from Trenton to Toronto delivering a donor organ for transplantation. The mission was time critical: crews were briefed that they had approximately one hour to arrive in Toronto for the organ to be viable. At this point they received a weather and operations briefing.

24. **The CC-130 Flight Simulator Scenario.** The scenario used in this study was constructed with the cooperation and expertise of CC-130 trainers in 426 Squadron. The scenario was devised to test several aspects of the mental model, especially selected systems knowledge and resource management skills. Moreover, the scenario was constructed to have significant training value for the participating aircrew. Thus, the test scenario was able to substitute for a standard simulator session, thereby minimizing the disruption to normal CC-130 simulator training. Efforts were made to make the simulator session as realistic as possible through the use of a five minute videotaped mission and weather brief employing actual operations personnel from CFB Trenton.

25. The flight was a winter, poor weather, night IFR (Instrument Flight Rules) mission. The weather brief indicated that there were few alternates available, including Trenton where the weather was expected to close in at or soon after departure. The emergency medevac was necessary because the major highway running through the area (the 401) was closed due to poor weather and the 436 SAR crew was on another mission.

26. During the mission a number of aircraft system failures were simulated and there were changes to ATC procedures and weather that required replanning (see Table 1). While it is clear that the scenario was busy, care was taken to ensure that it did not present an unrealistic level of workload for most crews. All 23 crews completed the simulator flight, albeit with varying degrees of difficulty. The systems malfunctions were expected to take up a great deal of the crews' attention. In fact, a critical measure was the ability of the AC to continue to monitor and assess more discretionary or less pressing aspects of the flight such as the mission status and the weather.

27. **Expert Rating Assessments of Highly and Less Proficient Crews.** We adapted the Aircrew Observation and Evaluation Scale (5) as the metric used by our SMEs to evaluate the performance of crews. Three experienced pilots (one civilian, and two military ICPs) provided proficiency rankings for each of the 23

crews. These were based upon independent multidimensional assessments of each videotape (e.g., assessments of safety concerns, decision making, and workload management).

Table 1. List of system malfunctions, flightpath and weather updates in the CC-130 flight simulator scenario.

Aircraft system malfunction 1:	On take-off landing gear will not retract (touchdown relay failure)
Aircraft system malfunction 2:	#2 EDHP light (pump malfunction)
Toronto ATC malfunction:	Toronto radar goes down, Toronto is on procedural control, the aircraft is directed to Simcoe to hold
Aircraft System malfunction 3:	#4 Generator light (generator failure)
Aircraft System malfunction 4:	#4 Generator bearing light (bearing failure)
Approaching Toronto wind updates:	Wind on arrival runway 24 at YYZ approaches crosswind limits
Aircraft system malfunction 5:	#1 reduction gearbox failure

28. Preliminary analysis of the crew communication data obtained from the videotapes indicated that crew effects were largely driven by the behaviour of the AC. This would be expected in hierarchically structured teams such as flightdeck aircrew. Indeed it was the ACs who made the majority of statements throughout the flights. A highly-proficient group of six crews and a less-proficient group of six crews were selected based upon the three rater's assessments of AC proficiency.

29. The three raters met as a group and reviewed their relative scorings for the 23 crews. The crews that all three raters had independently selected as representative of the higher and lesser proficiency groups were automatically included in our test group of crews. Finally the three raters debated their assessments of remaining crews to achieve consensus regarding the crews that were to be included in the highly and less proficient groups. Thus, it was only necessary to yield two groups which the raters agreed on average represented a more proficient group and a less proficient group. The proficiency groups included 12 ACs (6 highly and 6 less proficient), 12 CPs and 10 FEs (5 in the high

and 5 in the low proficiency groups)¹. As one might expect, a statistical test (*t* test) revealed that the highly proficient ACs had a greater number of hours on crewed aircraft (High: mean hrs. = 3800.83, Low: mean hrs. = 1819.17, *t* = 2.35, *p* = 0.05) and had spent more hours as ACs (High: mean hrs. = 2425.0; Low: mean hrs. = 883.33, *t* = 1.88, *p* = 0.11), than did ACs in the less proficient group, although this latter result is only marginally statistically significant.

30. Crew member communications from each of the twelve video tapes were coded by two independent coders (who were different from the raters of 'proficiency' and who were 'blind' to the proficiency group assignment of the crews) according to the mental model categories outlined in Appendix 1 of Annex B.

Analysis

31. The specific unit of analysis used here was the number of communications in each coding category made by a crewmember, divided by the total number of communications made by that crewmember. This provides a measure of the proportion of communications that fell within each of our coding categories. Essentially we asked the questions: "...Out of all the communications (statements, commands, questions etc.) made by an individual, what proportion of statements reflect each of our categories?" and, more importantly, "...Does the pattern of these communications reliably differentiate highly from less proficient ACs?"

32. On the basis of past theory and research, we expected that relative to lower proficiency ACs, highly proficient aircraft commanders would

- show a greater range of thought (think about more domains)
- show a greater depth of thought (think at deeper levels)
- show superior systems knowledge and resource management skills, and
- demonstrate fewer instances of open loop communications.

To make this determination, we conducted a series of one-way Analysis of Variance (ANOVA) analyses on the AC's data to determine those mental model variables that distinguish highly from less proficient aircraft commanders.

Results

33. **Highly Proficient ACs.** As the results presented in Table 2 indicate, the overall pattern of results substantiated our hypotheses. Specifically, highly proficient aircraft commanders (relative to less proficient aircraft commanders) demonstrated greater depth of thought, as evidenced by a greater level of

¹ In two sessions 426 flight engineer instructors who were naive to the experimental design and hypotheses and the details of the simulator scenario stood in for missing squadron FEs.

preplanning during the simulator flight ($t = 1.73, p = 0.05$), especially concerning procedures and checklists ($t = 2.05, p = 0.04$), geography or air picture ($t = 2.35, p = 0.02$), the weather ($t = 1.55, p = 0.07$), and the mission itself ($t = 1.26, p = 0.12$). Moreover, highly proficient ACs were also more likely to note the implications of changes in wind direction as they approached Toronto ($t = 1.76, p = 0.05$).

34. Also as anticipated, highly proficient ACs demonstrated a greater range of thought. Their statements encompassed a greater number of the mental model domains relevant to this flight scenario, but most particularly concerning the mission (awareness: $t = 1.40, p = 0.09$, total proportion of statements concerning the mission: $t = 1.45, p = 0.06$) and the weather (awareness: $t = 1.56, p = 0.08$, total proportion of statements concerning the weather: $t = 1.77, p = 0.06$). These results are particularly striking as they reflect the fact that highly proficient ACs were better able to keep in mind these more discretionary portions of the total flight mental model.

Table 2. Pattern of results of mental model domains and functions among highly and less proficient aircraft commanders.

		← R A N G E →				
(less)	Mental Model Content Domain	Mission	Geography (Air Picture)	Systems	Procedures & Checklists	Weather
D	Mental Model Function					
E	Awareness	H > L			H > L	H > L
P	Cross-Checking		H > L		L > H	
T	Implications					H > L
H	Preplanning	H > L	H > L		H > L	H > L
(greater)						

Note: only statistically significant or marginally significant results are reported here.

H > L = Highly proficient ACs make a greater proportion of these statements relative to less proficient ACs.

L > H = Less proficient ACs make a greater proportion of these statements relative to highly proficient ACs.

35. **Less Proficient ACs.** Our results also indicated that less proficient ACs engaged in greater cross-checking in terms of checklists/procedures ($t = 2.57, p = 0.01$). At first glance this may seem contrary to our hypotheses. However, this result simply reflects the fact that less proficient ACs were more unsure of the relevant aircraft systems and related checklists.

36. This lesser system knowledge on the part of less proficient ACs is further substantiated by the fact that they demonstrated less system knowledge according to our coding scheme ($t = 2.40, p = 0.02$ — see Table 3). Thus, the less proficient ACs were less able to spontaneously address system malfunctions without referring to checklist and manuals; they simply had less system knowledge at their fingertips.

Table 3. Pattern of results for additional coding categories.

SYSTEMS KNOWLEDGE	H > L
CREW MONITORING	H > L
TASK PRIORITIZATION	H > L
OPEN LOOP COMMUNICATION	L > H

H > L = Highly proficient ACs make a greater proportion of these statements relative to less proficient ACs.

L > H = Less proficient ACs make a greater proportion of these statements relative to highly proficient ACs.

37. Other results concerning resource management skills also differentiated the highly from the less proficient ACs. As Table 3 also indicates, highly proficient ACs showed some evidence of greater task prioritization ($t = 1.34, p = 0.10$). Moreover, these individuals were more likely to engage in crew monitoring, that is, be aware and concerned about the workload and stress levels of their crews ($t = 1.41, p = 0.09$). Finally, Table 3 also indicates that less proficient ACs showed a greater proportion of open loop communications than did highly proficient ACs ($t = 2.65, p = 0.01$). Indeed, less proficient ACs evidenced three times the instances of open loop communication, suggesting, as expected, less efficient information exchange at the crew level and a higher level of information or work overload for the less proficient ACs.

38. **Compensatory Behaviours.** We also analyzed the communications of the other crewmembers of the highly and less proficient ACs (i.e., CPs and FEs).

Overall, we saw an interesting pattern of behaviours emerge for the crews of the less proficient ACs. Specifically, the copilots of less proficient ACs tended to make more awareness statements ($t = 1.82, p = 0.05$), and attempted to take a more directive ($t = 1.63, p = 0.07$) role concerning systems malfunctions and issues. Furthermore, the copilots of the less proficient ACs made more awareness statements ($t = 2.56, p = 0.01$), and took a more proactive role ($t = 2.14, p = 0.03$) regarding checklists and procedures. We saw a similar pattern of results for the flight engineers of the less proficient ACs. These FEs tended to take a more proactive role concerning systems malfunctions ($t = 1.56, p = 0.07$), and to engage in more preplanning concerning checklists and procedures ($t = 1.51, p = 0.08$). Perhaps most descriptive of communication problems, there were greater instances of open loop communication among the crews of the less proficient ACs (CPs: $t = 2.70, p = 0.02$; FEs: $t = 1.58, p = 0.08$).

Discussion

39. The original aim of the CBAG was to develop a measurement battery of communication and behavioural patterns to be used in future research. We believe that we have achieved this goal. The measures developed were designed to capture known and hypothesized differences in proficient flight behaviours. Their validity was demonstrated in a known-groups design (i.e., a high versus less proficient AC comparison). The measurement battery presented here is reliable, capable of yielding scientifically defensible results based on theory, and is applicable to a wide range of operational issues of concern to ATG.

40. Beyond the development of a measurement battery, the methodology we selected also allowed us to begin to investigate important differences between highly and less proficient aircrew behaviour. Such distinctions are particularly relevant to training, that is, identifying those positive behaviours that should be particularly highlighted and modeled in the training system, as well as those specific behaviours that contribute to ineffectiveness and less safe practices among aircrew. Specifically, our results indicated that highly proficient ACs were characterized in terms of strong systems and procedural knowledge. Furthermore, they were more likely to demonstrate a superior range and depth of thought concerning important aspects of the flight and were more able to address discretionary aspects such as weather and the mission. Highly proficient ACs also demonstrated greater resource management skills at both the team level and at the level of their own activities, as evidenced by their greater awareness of the stress and workload levels of their crews and their greater tendency to prioritize tasks. Thus, our proficient ACs facilitated teamwork because their communications maximized the planning of flight-related tasks and goals. We have interpreted this as a result of their forming and transmitting to aircrew an effective and efficient shared mental model of the flight.

41. On the other hand, we found that less proficient ACs had less knowledge of the CC-130 systems. The impact of this lesser systems knowledge is that it

likely increased the overall workload of the AC and the rest of the crew. The less proficient ACs showed evidence of work or information overload as they were less likely to respond to the questions and statements of their crews (an indirect indicator of workload). Just as importantly, less proficient ACs also demonstrated less range and depth of thought. That is, they engaged in less preplanning than did their more proficient counterparts. Moreover, less proficient ACs seemed to focus primarily upon systems-related, procedural cross-checking and rechecking of information. Indeed, it may be that the lack of systems knowledge simply saturated the mental capabilities of the less proficient group of ACs and they simply did not have additional mental resources to address aspects of the flight that were more discretionary (a restricted opportunity hypothesis). Alternatively, these findings may also reveal that those ACs deemed less proficient simply do not typically demonstrate great range or depth of thought (a restricted capacity hypothesis). The present design does not determine whether restricted opportunity or restricted capacity is at the root of these findings.

42. With respect to the theoretical foundations of our work, our results indicate that more proficient ACs have better articulated mental models as evidenced by their ability to quickly identify and correct aircraft system errors, to understand the consequences and implications of flight anomalies and to preplan the remaining portion of the flight in light of these implications. In effect, high proficiency ACs are better dynamic decision makers and better purveyors of information to their crews. Integrating these results with DCIEM's past theoretical work suggests that highly proficient ACs with a sound systems knowledge base use preplanning to both decrease the time taken to process decision-relevant information and increase the time available to devise, coordinate, and action plans.

43. Our results also begin to illustrate the interactive and systemic or dynamic nature of crewwork. We found some evidence that the crews of the less proficient ACs attempted to compensate for the particular weaknesses of their ACs. Importantly however, these attempted behaviours did not fully compensate for the deficits of the less proficient ACs.

Implications of Findings to ATG

44. It is not possible to link the findings of this study to past CC-130 accidents in a way that establishes causality. That said, it can be reasonably inferred that the human factors characteristics of less proficient crews are related to unsafe flight, and hence, higher accident potential.

45. Highly proficient ACs had a greater number of flight hours on crewed aircraft and had more hours as aircraft commanders. With active airline recruiting of ATG pilots, experience levels will further decrease. To compensate for decreased experience levels, greater emphasis on training is needed to ensure

that all aircrew share the characteristics of the highly proficient crews. Past training programs were designed with the assurances of highly experienced ACs and continued tutelage of new pilots. A revised training system designed in light of current realities is indicated.

RECOMMENDATIONS

45. The following recommendations are made by the study group, based on the work of the CBAG, FPAG and collateral studies:

- a. Raise the overall level of systems knowledge among ATG aircrew by developing teaching aids and courses to accelerate the acquisition of knowledge and to compensate for lowered fleet experience levels through advances in training program delivery.
- b. Review the Aircrew Coordination Training (ACT) syllabus to ensure that its objectives are appropriately performance-based and emphasize the management of task load and the provision of timely decision making.
- c. Augment the ATG ACT with experiential decision making and leadership training, especially for ACs, via the development of flight simulator scenarios that test dynamic decision making, risk assessment and resource management skills.
- d. Conduct a feasibility study on the suitability of using the technology developed by the FPAG in analyzing flight recorder data. The study will focus on the ability to identify trends in the degree and frequency of unsafe flight conditions. (see the report of the FPAG at Annex A)
- e. Continue fatigue studies specific to long range strategic airlift (6).
- f. Explore the use of modern eye-tracking technology as an aid to instructional training (6).

REFERENCES

1. AIRCOM DCOMD 001 072200Z Jan 94.
2. ATG/DCIEM Human Factors Study of CC 130 Operations: Feasibility and Progress 20 September 1994 (Annex C).
3. 1630-2 (HPSD) 15 March 1995: ATG/DCIEM Human Factors Study of CC 130 Operations Progress Report (Annex D).
4. Report on Box Top study (Annex E).
5. Clothier, C. C. (1991b). Behavioral interactions across various aircraft types: Results of systematic observations of line observations and simulations. In *Proceedings of the Sixth International Symposium on Aviation Psychology*, 332-337.
6. Briefing to Cmdr ATG 16 July 1996, CFB Trenton, Canada.

ANNEX A. FLIGHT PERFORMANCE ASSESSMENT

Introduction

1. The FPAG worked as a sub-unit of the overall study group. The tasks of this group were: 1) to develop techniques for the analysis of the accuracy and consistency of the aircraft flight path; and 2) to determine if the accuracy and consistency of flight path performance correlated with the proficiency ratings used by the CBAG to assign crews to the highly and less proficient categories. The group was blinded to the work of the CBAG so that an unbiased comparison of the two methods could be made. The primary activities of the FPAG consisted of literature review, recording of simulator flight data, format translation, reconstruction of the flight path in order to identify specific phases of the flight, metric development, and scoring and analysis of simulator flights. As the method matured over the course of the study, the feasibility of using these metrics for the operational monitoring of CC-130 aircraft was also considered.

Literature Review

2. A review of existing techniques for using the aircraft state as an objective metric of crew performance was undertaken to: 1) identify an approach suitable for the analysis of performance on the CC-130 simulator and aircraft; 2) identify the statistical techniques to be used in the analysis of these parameters within a single flight; and 3) develop techniques to analyze composite trends in flight performance over many flights. A review of the literature revealed that there was no common approach to this problem. Also, there has been little formal research into the development of robust metrics of aircraft state that are sensitive to aircrew skill levels.

Simulator Data Processing

3. Data on each of the 23 simulated flights from the CBAG study were downloaded to tape and transported to DCIEM for analysis. Prior to these flights, simulator data were collected and used as a basis of developing the required technology. There were a number of difficulties in collecting and analyzing data from the CC-130 simulator due to the lack of documentation on the data storage formats. Once the data files were decoded, the flight path of each crew was reconstructed and compared to the nominal flight path of the simulated mission. The limited computer storage of the CC-130 simulator and subsequent loss of data limited the number of crews for which performance could be evaluated. As a result of these problems flight path data for only 15 out of 23 crews were available for analysis.

Metric Development

4. A commonly used method of evaluating performance is a simple binary

score, i.e., satisfactory or unsatisfactory. This approach is currently used by ATG during routine check rides to score broad areas of general airmanship and aircraft handling skills. Alternative metrics of flying performance include: maximum deviation from a nominal value, variability about the nominal value, and time-off-target, along with other combined measures of accuracy and variability which may reflect degradation in performance. However, these measures of flight path accuracy and variability do not allow for the specific identification of deviations from nominal values of the flight parameters in a way that indicates that an unsafe condition has developed. They were therefore not used in this study.

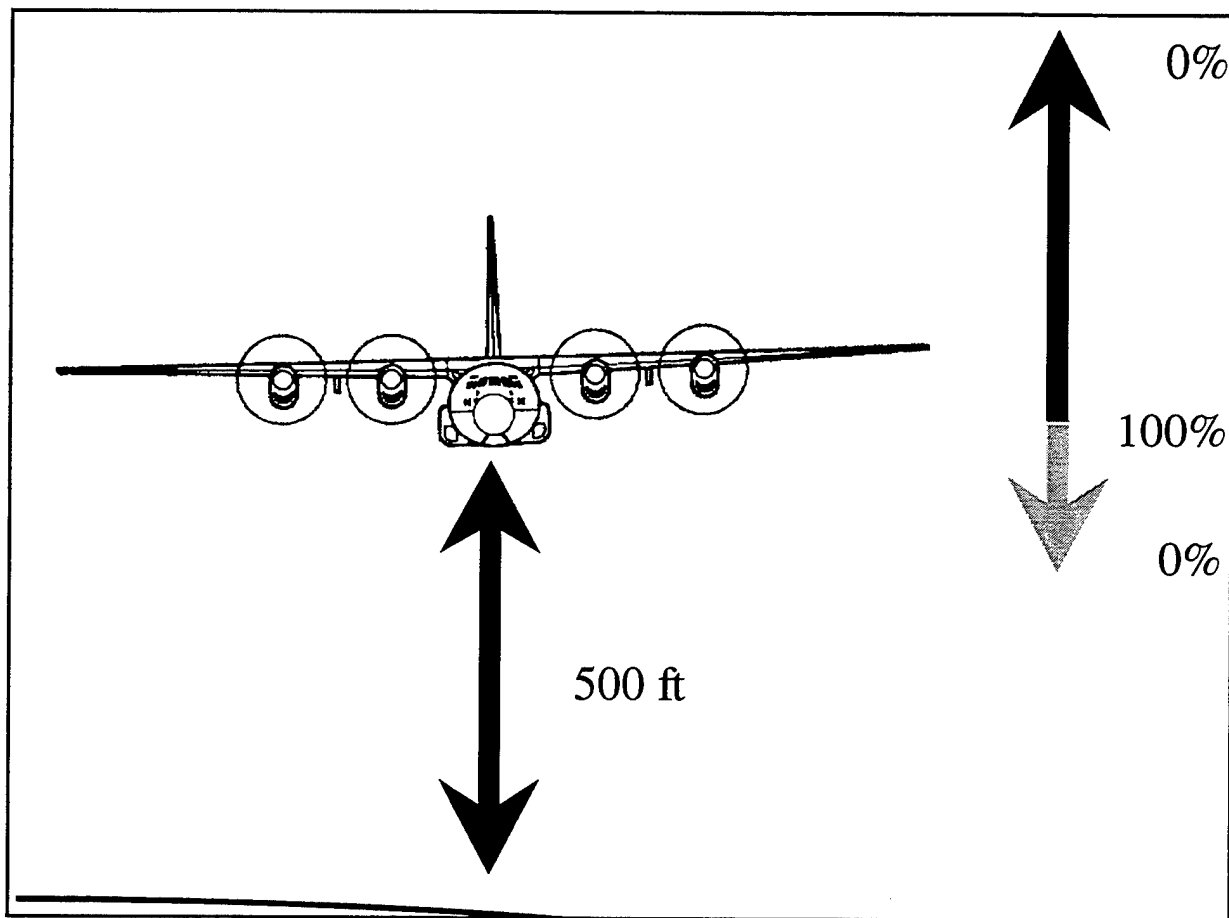


Figure A1. Example of the flight scoring method — illustrating deviations above and below the assigned altitude for low level flight over terrain such there is a significant probability of an accident.

5. A scoring metric based on the expertise of ATG standards and training officers was developed using a questionnaire to quantify how instructors would 'score' the performance of CC-130 crews during various stages of flight. The

questionnaire was presented in two parts. The first part served to rank, on a scale of 1 to 10, the importance of a set of flying parameters, glide slope, altitude, airspeed etc., for each phase of a CC-130 operational flight. Scoring metrics were generated for the takeoff, climb, strategic enroute, tactical enroute, descent, approach, and landing phases of flight. The scoring system was developed for a number of flight parameters including airspeed, altitude, bank angle, deviation from track, deviation from glideslope, takeoff speed, etc.

6. The second part of the questionnaire asked the standards officers to develop a performance function for the five most important flight parameters for each stage of the flight. For each parameter the standards officers were asked to assign values to the absolute deviations from the nominal values of the flight parameters which, in their opinion, would result in a significant probability of an accident.

7. Nominal flight performance with respect to a specific flight parameter was assigned a score of 100%. Any deviation from the nominal value of the flight parameter, which would be expected to result in a significant probability of an accident was scored as 0%. A performance function was generated for each parameter which identified the score for a given degree and direction of deviation from the nominal value of the parameter. An example of this is shown in Figure A1. If a CC-130 aircraft was flying at 500 ft during a SAR mission, a 100 ft deviation below that altitude might result in a score of 0% from a particular standards officer as a result of the risk of ground impact. However, a 500 ft deviation above the assigned altitude might be required before the score was 0% due to the risk of collision with another aircraft.

8. Twenty-nine standards and training officers in ATG completed the survey. The parameter rankings and the scoring functions from each survey were transferred to a statistical analysis software package at DCIEM. The scoring functions developed by each standards officer, for each flight parameter and segment of flight were averaged and used to analyze the data collected from the simulator study. Analysis of the CC-130 simulator mission focused on the enroute cruise and the landing segments of flight.

Results

9. The flight path accuracy and consistency analysis resulted in a clear cut stratification of crews with respect to the important flight parameters in three phases of the simulated flight, as well as a percentage score for each flight parameter and each phase of the flight. The crews were ranked with respect to their ability to maintain the assigned altitude and heading during the enroute phase of the flight and to maintain the correct glide slope during the landing phase.

10. The crew rankings generated from the survey based scoring functions

were compared to CBAG analysis, specifically the proficiency rankings generated by a civilian ICP and the rankings based on the 'mental model' scores (see para 16. main report). Spearman rank correlation coefficients were calculated for each comparison. There was a significant correlation between the rankings generated from the glide slope deviation scores and the proficiency rankings generated by the civilian pilot ($r = 0.55$, $p < 0.04$). Crew rankings derived from the deviation from altitude and deviation from heading scores correlated with the proficiency ratings at $p < 0.1$. There was a statistically significant correlation ($r = 0.85$) at the $p < 0.1$ level between the rankings generated from the deviation from heading scores and the mental model scores calculated by the CBAG.

Discussion

11. The weak positive correlation seen when comparing crew stratification based on assessment of flight path and crew behaviour suggests that crews who work well together are most successful in flying the aircraft efficiently and safely. This supports the use of proficiency scores by the CBAG to categorize high and low proficiency crews.

12. Early in the study, it was determined that this technology could be used on actual flight decks as a statistical monitor of events. Programmed to detect and assess high risk excursions from the data stored on the flight data recorder, this technology could provide ATG with statistical analysis of aircrew performance. A number of European airlines currently monitor the performance of aircrew, providing feedback when significant exceedences from nominal flight path parameters are detected. NASA Ames is undertaking a study into the development of an automated performance monitoring system for US airlines.

FPAG Recommendation

13. We recommend that the technology developed by the FPAG be evaluated for use in analyzing flight recorder data, focusing on the feasibility of using this technology to identify trends in the degree and frequency of unsafe flight conditions in the CC-130. This evaluation would entail:

- a. collecting data from a number of CC-130 H model flights in order to develop the data extraction and analysis algorithms;
- b. revising, enhancing, and implementing the ATG/DCIEM aircrew performance metrics survey as a computer program;
- c. developing enhanced flight path analysis metrics based on the revised survey and analysis techniques; and
- d. developing software to track trends in flight path deviations.

ANNEX B. THEORETICAL BASIS FOR ASSESSING CREW BEHAVIOUR

1. The Information Processing/Perceptual Control Theory (IP/PCT) Model (1) provided a theoretical basis for the work presented in this report. A major assumption of the model is that poor performance stems from processing overloads which causes information to be shed. Such overloads occur when time pressures become excessive. The model states that operator workload, errors and performance are driven by the following time pressure ratio:

$$\frac{\text{time required to process the information necessary to make a decision}}{\text{time available before the decision must be actioned}}$$

2. This ratio represents both real and felt time pressure as the perceived time pressure, which will determine human response to the imposed load, depends on the perceived time available. When the time to process decision-relevant information exceeds the time available to make the decision, resulting in a time pressure ratio greater than 1, information remains unprocessed and is either shed or may be stored in memory for later processing and implementation. In this case the individual should take actions to reduce the resultant time pressure by either: i) reducing the time required to process the information (via information or task shedding, delegation to another crewmember, or sometimes, by making a less accurate decision), or ii) buying additional time to make the decision, or iii) through some combination of i) and ii). Effective decision-making reduces time pressure and excessive workload on the flightdeck. This is facilitated by a variety of factors, for example:

- a. superior mental models (see the section below) and knowledge of aircraft systems and procedures (i.e., expertise) reduce decision times;
- b. a large repertoire of related situations in which similar decisions have been made before (i.e., experience) also speeds decision making;
- c. the ability to note deviations from planned flight status (i.e., situation awareness, awareness updates and cross-checking) maintains accurate mental models and reduces uncertainty;
- d. the ability to draw implications from changes in flight status, and the processing of information that relates to contingency planning, allow workload peaks to be smoothed by the proactive resolution of uncertainty; and
- e. superior crew management and coordination skills that creates a working environment where rapid information processing and sharing can occur.

Mental Models

3. As noted in the section above, a key component of efficient decision making is the quality of the mental model that an individual holds. Hendy (1) has defined mental models as:

“...that part of the operator’s internal state which contains the knowledge and structure necessary to perform a task. As such the operator’s mental model directly shapes the operator’s responses and determines the potential to perform in accordance with system or task demands. The mental model contains the operator’s goal state and provides the reference against which actions are selected and initiated...”

4. In essence then, a mental model is ‘the what, who, why, and the how of a task’. A mental model may encompass past, present, and future flight parameters, goals, and considerations. Recall that well developed mental models lead to more efficient information processing, to decreased time pressure and workload, and to better performance.

5. Mental models can be conceptualized in terms of the domain they address (the what) and in terms of the function they serve (the how). A mental model **domain** refers to content: the discernible and distinguishable aspects or domains of an individual’s thoughts concerning a flight. Theoretically, there can be any number of mental model domains, some being more or less relevant to a given flight scenario. In the present scenario the following domains are central: aircraft systems, procedures and checklists, geography or air picture, the weather, and the mission itself.

6. The **function** of a mental model refers essentially to the process involved in performing a task and may be manifested in the verbal communications among the crew. The goal of a communication can include alerting or awareness (referring to simple awareness regarding the state of some domain of the mental model); cross-checking (the explicit noting of deviations with respect to another crew member’s behaviours, planning, calculations, decisions etc.); noting implications (the explicit and overt recognition of the consequences or ramifications of present states and the interactive effects of these states); and preplanning (explicit statements of future intentions, behaviours or plans regarding any mental model domain).

7. Mental model domains and functions can, of course, be considered together (see Table B1). Moreover, one can think of domain and function as reflecting the range and the depth of the mental model respectively. For instance, the greater the number of mental model domains considered during a flight, the greater the range of thought demonstrated by an individual. Similarly, simple awareness statements, for example, noticing a system

malfunction, would be classified as requiring less depth of thought than would noting the implications of system malfunctions or preplanning in light of these implications.

Table B1. Mental model domains and functions.

		<----- R A N G E ----->				
(less) D E P T H (greater)	Mental Model Content Domain ----- Mental Model Function	Mission	Geography (Air Picture)	Systems	Procedures & Checklists	Weather
	Awareness					
	Cross-Checking					
	Implications					
	Preplanning					

Shared Mental Models

8. Mental models become even more complicated when a task is to be completed by a team of individuals. That is, overlaid upon the individual's view of the task and how best to accomplish it is his or her preconceptions of how teams should function as well as the more specific conceptualization of how this particular team will function. Moreover, individual mental models about a task held by a particular crewmember must complement each other in order to provide a common basis of understanding which will allow effective coordination of behaviours. Researchers in the area suggest that it is the quality of the shared mental models that is chiefly responsible for the consequent effectiveness or the lack of effectiveness of a team (2,3,4).

Coding Mental Model Categories

9. Appendix 1 to this Annex outlines the coding scheme for the mental model categories used in the CC-130 study. This Appendix lists the categories by domain and function, and provides examples of the types of communications that would fit each descriptor.

References

1. Hendy, K. C. (1994). Implementation of a human information processing model for a task simulation network. (DCIEM Report Number 94-40). Defence and Civil Institute of Environmental Medicine, North York ON.
2. Ginnette, R. C. (1993). Crews as groups: Their formation and leadership. In B. G. Kanki, E. L. Weiner, & R. L. Helmreich (Eds.), *Cockpit Resource Management*. Orlando, Florida: Academic Press, 70-98.
3. Cannon-Bowers, J., A., Salas, E., & Converse, S. (1993). Shared mental models in expert team decision making. In J. Castellan (Ed.), *Current Issues in Individual and Group Decision Making*. Hillsdale, NJ: Erlbaum, 221-246.
4. McIntyre, R. M., Morgan, B. B, Salas, E., & Glickman, A. S. (1988). Teamwork from team training: New evidence for the development of teamwork skills during operational training. In *Proceedings of the Interservice/Industry Training Systems Conference*, Orlando, Florida.

ANNEX B

Appendix 1. Communication and Crew Behaviour Coding

NOTE: A COMMUNICATION MAY BE CODED WITH MORE THAN ONE MENTAL MODEL DOMAINS AND FUNCTION

DOMAIN

1. Mission

CONTENT DOMAIN OF COMMUNICATION	CODE	DESCRIPTION	EXAMPLES
MISSION	M	any mission specific communication (statement, question, command, etc.) relating to the time-limited, medical purpose of the flight	"We will ERO the organ through the cargo door." (AC-medic) "If you have any requirements or requests you can relay them to me through the loadmaster once we are enroute. Is there anything special you want us to take care of right now?"

2. Air Picture (Geography)

CONTENT DOMAIN OF COMMUNICATION	CODE	DESCRIPTION	EXAMPLES
AIR PICTURE (GEOGRAPHY)	G	any communication (statement, question, command, etc.) relating to the aircraft's physical place in space, including airways	"Plan on 24 left to Toronto." "From Trenton to Toronto the VOR is 36 miles."

3. Systems

CONTENT DOMAIN OF COMMUNICATION	CODE	DESCRIPTION	EXAMPLES
SYSTEMS	S	any communication relating to any of the mechanical, electrical, or navigational systems on board the aircraft. That is, any utterance that refers to the adjusting of a system will be classified as system-related, even if that system is geographical in nature (see example).	"For departure brief right seat let's start with Trenton on the TACAN and we'll check the localizers and switch them over to Campbellford before we start rolling." "Just tell me when you want the DME on Campbellford."

4. Weather

CONTENT DOMAIN OF COMMUNICATION	CODE	DESCRIPTION	EXAMPLES
WEATHER	W	any communication relating to weather e.g., wind direction, gusts, icing conditions.	"Okay PMA, we'll do a flap 50 because of the winds..." The winds are going to keep coming from the north. Must be a front moving through."

5. Procedures and Checklists

CONTENT DOMAIN OF COMMUNICATION	CODE	DESCRIPTION	EXAMPLES
PROCEDURES and CHECKLISTS	CHLST	any communication referring to checklists actioned or to be actioned, specific procedures, also any communications referring to clearances	<p>"OK, whenever you are ready, you should probably do a checklist here."</p> <p>"OK when you are ready right seat we'll ask them for start clearance."</p> <p>"We will do a PMA to Toronto."</p> <p>"Entering the procedure turn." [Note: This is also geographical and it indicates knowledge of where they are in space.]</p>

FUNCTION

1. Awareness

FUNCTION OR PURPOSE OF COMMUNICATION	CODE	DEFINITION	EXAMPLES
AWARENESS	AW	<p>simple declarative statements concerning the present states, and immediately impending* changes in the states of systems, the mission, air picture, weather, checklists. These statements serve to maintain and update the changing mental model and short-term goals.</p> <p>*as distinguished from preplanning</p> <p>information or opinions provided by crewmembers in response to a query or in the course of a discussion about systems, Wx, the mission, chscklists, or geography</p>	<p>"I'll slow it down a bit there."</p> <p>"Leveling 6,000 in the turn."</p> <p>"OK 8 miles back from Campbellford in the TACAN."</p> <p>"We've got a hydraulic low pressure light on #2."</p> <p>"OK, I'm going to push the power up."</p> <p>"Gear is up."</p> <p>"[ATIS is] coming up on Victor 2 right now."</p>

2. Cross-Checking

FUNCTION OR PURPOSE OF COMMUNICATION	CODE	DEFINITION	EXAMPLES
CROSS-CHECKING	CC	explicit, overt noting of deviations with respect to another crewmembers actions, planning, calculations etc.	<p>"You're off the glide slope."</p> <p>"Watch your airspeed."</p> <p>"You probably want to keep your airspeed down."</p> <p>"You're kind of diverging off the west from the VOR ..."</p> <p>"OK, if you want to fight her back down to 4 and show me up, that would be good."</p>

3. Implications

FUNCTION OR PURPOSE OF COMMUNICATION	CODE	DEFINITION	EXAMPLES
IMPLICATIONS	IMP	explicit and overt recognition of the consequences or ramifications of present system, mission, weather, checklists, and geographical states and the interactive effects of these states upon each other.	<p>"Just considering your worst case scenario, would you be interested in lowering your gear and flaps into position now, just in case the other pump goes?"</p> <p>"The only thing I've got doubts about is the antiskid system because it is part of the touchdown relay."</p> <p>"... other systems lost, we have no utility so we want to make sure that the brakes go to emergency."</p>

4. Preplanning

FUNCTION OR PURPOSE OF COMMUNICATION	CODE	DEFINITION	EXAMPLES
PREPLANNING	PP	explicit statements of future intentions or plans regarding systems, the mission, air picture, checklists & procedures	<p>"Okay, so we are 15 miles back from Simcoe and if we do a teardrop so that we're back on the 047 ... from there we do a procedure to get us into 24R."</p> <p>"350 a good trot to Campbellford till we're airborne and we'll pick it up and hold it. 257 out of Campbellford to TO."</p> <p>"Depending on RWY conditions, I'd go to emergency or deactivate your antiskid just to be safe and basically be gentle on your brakes."</p> <p>"What you do is over the VOR outbound, turn right inbound and come in 255."</p> <p>"As soon as we come out of the hold do you want to call the approach check?"</p>

ADDITIONAL CODING CATEGORIES:

1. Systems Knowledge

ADDITIONAL CODING CATEGORY:	CODE	DEFINITION	EXAMPLES
SYSTEM KNOWLEDGE	SK	<u>spontaneous</u> & explicit statements by crewmembers regarding how relevant systems operate, indicating their knowledge of systems as distinguished from reading information in flight manuals (e.g., the Z-10).	[re: gear override] CO: "they're not coming up" AC: "Hit the override switch there."

2. Task Prioritization

RESOURCE MANAGEMENT	CODE	DEFINITION	EXAMPLES
TASK PRIORITIZATION	TSKPR	overt & explicit direction or request for direction concerning the timing of tasks in relation to one another	"Let's get the approach check done and then we'll deal with that ..." "I'll finish the post-takeoff check and then we'll deal with the landing gear problem before we get too far from Trenton."

3. Crew Monitoring

RESOURCE MANAGEMENT	CODE	DEFINITION	EXAMPLES
CREW MONITORING	CM	overt & explicit statements directed toward assessing and addressing crewmembers' workload, information load, or vigilance	"Just let me know how you are doing over there." "How are you doing there CO?" "Everybody got the ATIS?"

4. Open - Loop Communication

RESOURCE MANAGEMENT	CODE	DEFINITION	EXAMPLES
OPEN-LOOP COMMUNICATION	OLC	lack of response of any kind to directions, suggestions, queries; also interruptions that <i>interfere</i> with building a mental model, information transfer, or the development of a team, <u>as distinguished from alerting interruptions</u> (e.g., FE must break in to alert the AC to a system emergency), or interruptions that involve one crew member picking up on and finishing another's thoughts and statements (which is evidence of a shared mental model).	

ANNEX C. ATG/DCIEM HUMAN FACTORS STUDY OF CC 130 OPERATIONS: FEASIBILITY AND PROGRESS 20 SEPTEMBER 1994

(HPSD)

September 1994 (replaces 9 Sept 94 version)

Background

1. Prior to 1979, the Canadian Forces (CF) experienced only one fatal CC-130 air accident during 15 years of operations. After 1979, there were six fatal air accidents that involved seven aircraft. Although small numbers precluded statistically based conclusions, a perception grew that unsafe conditions existed since 1979 that contributed to an increase in the incidence of fatal air accidents. As a result of this perception, Air Transport Group (ATG) initiated studies on various aspects of the CC-130 operation to learn of any unknown factors that had contributed to a trend towards increased accidents. As well, AIRCOM requested that DCIEM undertake study of human factors issues in CC 130 operations (1,2,3). This request was made because of the success of DCIEM's survey of human factors in CF-18 operations (4).

2. In addressing the AIRCOM request, this report proposes a study aim, reviews progress to date, offers an approach, and speculates on the feasibility of completing the proposed study.

Proposed Aim

3. To establish the effect of selected human factors issues on flight safety in the CC 130, as determined by crew performance, and to recommend measures to prevent future accidents.

Progress To Date

4. It was initially suggested that this study would be similar in design and execution to the earlier DCIEM CF-18 survey of human factors. Following review by DCIEM staff, and discussion with ATG, opinion changed. It was concluded that many of the human factors identified as affecting CF-18 operations were also applicable to the CC-130 operation. Several of those factors were related to policy issues which were beyond the ability of ATG to change. For these reasons, the option of replicating the CF-18 human factors survey was considered, and rejected in favour of an approach that would build on the earlier study results. Rather than a survey of human issues, a detailed evaluation of select human factors issues relevant to the CC 130 was considered a superior approach.

5. During a search for an alternate approach, ATG conducted an internal

review and identified the following areas of primary concern (5):

- a. the primacy of operations superseding flight safety;
- b. unsafe operating procedures;
- c. low experience levels;
- d. eroded standards and training;
- e. incompatible lifestyles;
- f. insufficient guidance from higher headquarters.

6. In reviewing this list, DCIEM/ATG staff noted that each item on ATG's list of concerns almost certainly affects the process of conducting flight operations. Determining the relative effect of each of these concerns (and others), or interactions was identified as the problem. In conducting research that would address this list of concerns, the problem of creating a study design that would lead to scientifically defensible conclusions became the central issue.

7. Prior to resolving this issue, DCIEM completed a survey of six A category CC-130 accident boards in order to further assess the scope and requirements of the study.

Review of CC 130 Mishaps

8. Over the last 14 years the CF has experienced six A category CC-130 accidents. In each, there were fatalities and the aircraft was destroyed. Presently, ATG has a higher number of flight safety air incidents and accidents compared to even ten years ago, and much higher than that of other countries such as the Royal Air Force. All of these accidents were influenced or caused by human cause factors, as opposed to aircraft malfunction. Failures due to human factors are often the result of extremely complex interactions of the task, equipment and crew; however, in most of these accidents, the crew could have corrected the problem in time to prevent the severe consequences.

9. All of the reviewed accidents illustrated an element of crew procedure or coordination breakdown, as the following examples demonstrate:

- a. a SAR mission ended when the aircraft stalled due to inappropriate procedure by the pilot. It is thought that the stall would have been recoverable if the correct procedures had been used. Neither the co-pilot nor the FE took any action even though they had each noticed irregularities during the flight;
- b. a formation fly-past by three CC-130s led to a fatal mid-air collision

when the lead pilot did not fly the briefed profile; and

- c. a high Arctic crash occurred when the aircraft captain apparently misidentified his position and descended to an unsafe altitude. No other crew member alerted the pilot that he was below his briefed altitude, nor did any of the crew realize the danger of the situation.

10. The survey revealed that breakdown of crew coordination played a role in the majority of occurrences. Human factors directly effect crew coordination, which can be revealed through assessment of crew performance. ATG is currently addressing crew coordination with the Aircrew Coordination Training (ACT) program, but the effect of this program on flight safety is unknown. The ACT program was identified as an important human factors issue that should be given priority consideration.

Overall Approach

11. The purpose of a multiple crew cockpit design is to allow division of tasks in a manner that reduces workload and therefore reduces error that is caused by overtasking. The requirement for a multicrew system is generally based on workload demands. For a team to operate 'error free', the workload levels of each crew member must be acceptable (less than 70-80% time occupied) and each must share a common (or at least compatible) mental model. Another purpose is to allow crew members to be able to compensate for errors made by others. In that sense, a multiple crew is, ideally, a self-correcting entity that operates with very few errors. Thus, an approach that measures workload, or error detecting and correcting, may provide information which could be used in scientific study.

12. Scientifically valid conclusions are possible by measuring an output and comparing the resulting data for different conditions. Crew behaviour during flight operations is a "process" that can be observed. Movement of the aircraft through the atmosphere is a "process" that can be recorded through flight data systems and analyzed. What is required for the proposed study is a measure of the "output" of these processes which is related to flight safety and operational effectiveness. Accident statistics are the most appropriate "output" measure, but they are too few in number to use to draw valid conclusions. Another possible "output" measure, yet to be developed, would be based on metrics that score crew behaviour, and aircraft flight parameters, against an ideal.

13. At the July 1994 SMC and ATG briefings, we proposed that this study develop such metrics. We recommended that the metrics be state-of-the-art tools that measure two separate "processes:" crew behaviour, and flight performance.

Measurement of Crew Behaviour

14. The implicit goal of the crew behaviour part of the ATG project is to measure some aspect of aircrew behaviour which correlates with flight safety and operational effectiveness. The ideal dependent variable would be air accident events. However, CF operations are too small in scale to permit the number of observations required to provide "scientifically defensible conclusions" about flight safety. Ideally, what is required is an outcome measure reflecting flight performance and safety ("outcome measures" provide information about the result of a process, as distinct from "process measures"). This outcome measure must have a high enough occurrence rate (sample size) to give statistically valid data.
15. Since there is presently no outcome measure for crew performance that correlates with flight safety, an intermediate or process measure, must be developed. What is required is some measure of "safe" team behaviour which, if increased, will result in increased flight safety. Crew Resources Management (CRM) is an approach to training that has been associated with quite dramatic improvements in flight safety (6). In CRM training, team behaviour is evaluated using a number of process measures. No data have been found on the relative contribution to flight safety of each of these parameters, although one report has been found of correlation between mission effectiveness and CRM performance in a simulator (7).
16. Measurement scales used in CRM training are in an early stage of development (8). Although the theoretical basis for CRM training is grounded in observations of poor use of human resources (9) and the argument that the human sub-system in the cockpit must function effectively, the theoretical basis of the measurement scales is unclear.
17. To accomplish the aims of this study, the initial work must focus on identifying parameters of crew behaviour which can be related to flight safety and operational effectiveness. This will be done using two approaches.
18. One approach will review the literature on crew performance and safety, and will develop a theoretical basis for crew performance measures. Hendy (at DCIEM) is now in the process of reconciling the CRM approach with his information processing model based on the concept of managing attention resources of the human operator(s). This may provide the necessary theoretical framework for the development of crew behaviour measurements.
19. The other approach will seek to identify those characteristics which aircrew believe distinguish the performance of different crews. The most feasible approach would be to use the Personal Construct System of knowledge elicitation (10), although other knowledge elicitation techniques might be appropriate (11).

20. The two approaches will then be reconciled. Once a list of possible parameters have been identified they will be tested for consistency both within subject pools and between raters. The list of parameters, screened for reliability, must then be applied in widely differing situations, to determine the sensitivity of the scales to differences in crew behaviour.

21. In summary, a complete review of existing literature is required, but developmental work is probable. The aims of the study are considered achievable.

Measurement of Flight Performance

22. Air accidents involve significant excursions from planned flight parameters. This portion of the study assumes that significant excursions in planned flight parameters that do not result in an air accident nevertheless provide a measure of crew performance that correlates with flight safety and operational effectiveness. Assessment of crew flight performance would be accomplished through detailed collection and analysis of data deriving from simulator or aircraft flight data systems. These data would then be analyzed to provide a measure of crew flight performance. A visit to the CC-130 simulator found that it is highly probable that flight data and video recordings could be made at the facility with minimal difficulty. Discussions with the Simulator Support Officer and Simulator Technical Support staff determined that flight data from the simulator are recorded on a CDC 80 MByte disk drive by the simulator computer. These data can be transferred to 9 track 800/1600 bpi reel-to-reel magnetic tape at the simulator facility. DCIEM has a 9 track tape drive installed on the "dretor" computer of the central Sun facility which should be able to read the data files from the simulator.

23. At the present time only the last 20 minutes of any simulation session can be stored. Simulator staff made several suggestions for increasing the size of the stored data file. They include:

- a. a re-write of the computer software to increase the size of the disk sector used for data storage;
- b. re-installation of a second hard drive with additional software modifications. ATG simulator staff expressed interest in participating in the study.

24. The technical challenges can be solved. The aims of this portion of the CC 130 study are considered achievable.

Follow-on Studies

25. The end-point of the initial phase will be completion of development of two measurement tools: one that precisely measures crew behaviour; another that measures flight performance. These separate tools will be reconciled to provide differing measures of similar cockpit activities. Combining data from these separate measures should provide greater scientific precision, and allow accurate conclusions.

26. The developed tool will then be employed to research any human factors issue of interest to ATG. We have selected the current ATG Air Coordination Training (ACT) program for study. By utilizing the previously developed measuring tools, a study will be designed to examine the effectiveness of the ACT by comparing data from trained crews, with data from untrained crews.

27. Other human factors issues will then be identified for future study. This could include ATG's previously described list of concerns (2). If this study proves successful, ATG will have a powerful tool for studying human factors issues. This tool may be adaptable to other cockpit environments, and other Air Force groups.

Summary

28. AIRCOM requested that DCIEM undertake study of human factors issues in CC 130 operations. Initial work has established that human cause factors have contributed to an apparent increase in the incidence of fatal air accidents in the CC 130 operation. In studying human factor issues, it was determined that conclusions should be scientifically valid and defensible. It was proposed that the DCIEM/ATG CC-130 Human Factors Study develop metrics to facilitate scientifically defensible conclusions related to human factors issues in CC 130 operations. The metrics should be centered around crew behaviour, and flight performance. In developing a measure of crew behaviour, a theoretical framework for looking at resource management issues will be developed, followed by development of resource metrics within the theoretical framework. In developing a measure of flight performance, an initial assessment of the technological problems indicates probability of success. The aims of both aspects of the study are considered achievable. Following successful development of these metrics, a study will be designed to examine the effectiveness of the ACT by comparing data from trained crews, with data from untrained crews. A successful product could provide ATG with a powerful tool for studying human factors issues.

Recommended Plan

29. The study group will consist of DCIEM, ATG, AIRCOM and CRAD staff arranged into three separate, but interdependent groups (see Appendix 1). A detailed plan is included at Appendix 2. Estimated requirements are described at Appendix 3.

R.D. Banks
Deputy Director
Human Protective Systems Division
for Chief

References

1. AIRCOM 6681-8 (DComd) 3 Dec 93
2. AIRCOM DCOMD 001 072200Z Jan 94
3. 2900-1 (DCOMD) 23 Dec 93
4. DCIEM Report No. 91-11 Jan 91
5. 11500LT-1 (A/COMD) 28 April 1994
6. Alkov, R.A., (1994). Enhancing safety with aircrew coordination training, *Ergonomics in Design*, April, 13-18.
7. Prince, C., and Salas, E. (1993). Training and research for teamwork in the military aircrew, In: E.L. Weiner, B.G. Kanki & R.L. Helmreich (Eds.), *Cockpit resource management*, San Diego: Academic Press, 337-366.
8. Helmreich, R.L. & Foushee, H.C. (1993). Why Crew Resources Management? Empirical and theoretical bases of human factors training in aviation, In: E.L. Weiner, B.G. Kanki & R.L. Helmreich (Eds.), *Cockpit resource management*, San Diego: Academic Press, 3-46.
9. Ruffell-Smith, H.P. (1979). *A simulator study of pilot workload with errors, vigilance, and decisions*. NASA Technical Memorandum 78482. Moffett Field, CA: NASA-Ames Research Center.
10. Gaines, B.R. and Shaw, M. (1980). New directions in the analysis and interactive elicitation of personal construct systems, *Int. J. Man-Machine Studies*, 13, 81-116.

11. Godet, M., (1991). *Problemes & methodes de prospective: boite à outils*, Paris: Editions Futuribles GERPA.

ANNEX C

Appendix 1. Organization of the Study

(HPSD)

September 1994

Study Group

LCol Bob Banks - chairman (DCIEM)

LCol John Jensen - co-chair ATG (ATG)

LCol Mark Tysiaczny - AIRCOM rep

Mr. Bill Noble - CRAD rep

Maj Barry Davis - ATG Coordinator

Keith Hendy - Crew Behaviour Assessment

Bill Fraser - Flight performance assessment

Capt Helen Wright - CRM issues/central coordinator

Crew Behaviour Assessment Committee

Aim: To develop a method of precise measurement of crew behaviour in the CC 130.

Keith Hendy - Committee Head

Capt Helen Wright

One human factors specialist as primary responsibility

Other Human Factors Division/ ATG staff as required

Flight Performance Assessment Committee

Aim: To develop a method of precise measurement of flight performance in the CC 130.

Bill Fraser - Committee Head

One statistician/computer specialist as primary duty

ATG Operational Researcher

Other HPSD/ATG personnel as required

ANNEX C
Appendix 2. Work Plan

(HPSD)
September 1994

Work Plan Item

Evaluation of technical feasibility	12 Sep 94
Hire 2 consultants on contract	15 Oct 94
Literature review	15 Dec 94
Familiarization flights for study group members	1 Jan 95
Design of methodology	15 Mar 95
Evaluation of each method	15 May 95 Final
Design of methodology	15 Jul 95
Final report on development of experimental method	15 Oct 95
Conduct an experiment on current ATG CRM	15 Apr 96
Identify additional human factors for further eval	15 Apr 96

Reporting Plans

Initial feasibility report to SMC DCIEM and ATG	12 Sep 94
Progress report to DCIEM/ATG	15 Mar 95
Final report on method development	15 Oct 95
Final report on CRM experiment	15 Apr 96
Final report on additional human factors	15 Apr 96

Estimated Start Date: 15 Oct 94

Estimated Completion Date: 15 Apr 96

ANNEX C

Appendix 3. Financial and Personnel Requirements

(HPSD)

September 1994

Financial

TD funds for familiarization and liaison with other agencies and 2 full time workers for 1 year (contract)

Personnel

Study Group

4 part time

Committees

2 full time

6 part time

ANNEX D. 1630-2 (HPSD) 15 MARCH 1995: ATG/DCIEM HUMAN FACTORS STUDY OF CC 130 OPERATIONS PROGRESS REPORT

(HPSD)

March 1995

Reference: A. ATG/DCIEM Human Factors Study of CC 130 Operations:
Feasibility and Progress 20 September 1994.

Background

1. The objective of the joint ATG/DCIEM Human Factors Study is to provide a method for assessing the impact of various human factors variables on the safety and efficiency of Air Transport Group (ATG) CC-130 operations. Approval for the study was given by the Commander ATG and DCIEM Senior Management Committee (SMC) based on several months of preliminary work. The study commenced on 15 October 1994 in accordance with reference A.

2. The study is organized around two working groups: the Flight Performance Assessment Group; and the Crew Performance Assessment Group. Activities of these two working groups are supervised by a study group that includes the two study chairs, the head of each working group, and ATG, AIRCOM, CRAD and 10 TAG representatives.

Aim

3. This report describes the activities of the study group and each of the two working groups, and outlines the way ahead.

Progress to Date

4. In anticipation of ATG/DCIEM SMC approval, a Task Description Sheet (TDS) was sent to AIRCOM in September 1994 for sponsor approval of the proposed study. The TDS was approved by AIRCOM and CRAD by 19 September 1994, and funds were made available in October 1994.

5. The study group, including various members of the two working groups, has met at weekly intervals. Non-DCIEM members have attended meetings when able, and several meetings have been held at 8 Wing Trenton to accommodate ATG members. In response to interest from 10 TAG, an observer from 10 TAG has attended meetings. Notes of each meeting have been circulated to those unable to attend. An updated description of the organization is found at Appendix 1.

6. The current status of the work plan is:

<u>Task</u>	<u>Original Deadline</u>	<u>Status</u>
a. Evaluation of technical feasibility	12 Sep 94	Complete
b. Hire 2 consultants on contract	15 Oct 94	Complete Jan 95
c. Literature review	15 Dec 94	Ongoing
d. Familiarization flights for study group members	1 Jan 95	Ongoing
e. Final design methodology	15 Jul 95	15 Sep 95
f. Draft report on development of experiment method	15 Oct 95	15 Oct 95
g. Complete an experiment on current ATG CRM	15 Apr 96	15 May 96

7. As a result of delays in hiring of the consultants, some of the deadlines have been moved back several months. The preliminary literature review will be complete by the end of March 95. Initial familiarization flights will be completed by the end of April 95, but both activities will be ongoing throughout the project so a firm completion date has not been assigned. The status of the reporting plan is:

<u>Report</u>	<u>Original Deadline</u>	<u>Status</u>
a. Initial feasibility report to ATG and SMC/DCIEM	12 Sep 94	Complete
b. Progress report to ATG/DCIEM	15 Mar 95	Complete
c. Final report on method development	15 Oct 95	15 Dec 95
d. Final report on CRM experiment	15 Apr 96	15 Jul 96
e. Final report on additional human factors		15 Jul 96

Flight Performance Assessment Working Group

8. The Flight Performance Assessment Working Group is developing techniques for the collection, storage, processing and analysis of flight performance (e.g., attitude, altitude, airspeed, glide slope deviation, etc.) and cockpit video data from the CC-130 Hercules Simulator.

9. **Collection.** The simulator computer's hard disk can currently record only 20 minutes of flight data. However, multiple sessions can be saved onto magnetic tape, with only a few minutes of simulator down-time required to transfer flight data from disk to tape. It appears that the simulator's computer code can be modified and a second disk drive re-installed to allow for a more extended record to be stored and transferred to tape. Experiments in the simulator will require a video record of cockpit activities. A video capability has already been installed.

10. In-flight aircraft data collection will likely be feasible for future studies. Data from all models of CF CC-130 Hercules aircraft are transmitted throughout the aircraft on a common Harvard digital avionics bus. The four tanker aircraft in the fleet have a 24 hour flight data recorder (FDR) that stores over 60 flight parameters. This record can be downloaded to a portable computer. The other aircraft in the ATG inventory allow only 25 minutes of data storage in the FDR.

Assuming appropriate authorization, there appears to be no technical obstacle to installing DCIEM hardware in the aircraft that would tap into the existing avionics bus and record longer sections of flight data. Additional data will be collected from squadron sources (i.e., crew specific information, scheduling, etc.), through commercial sources, or through the Internet (i.e., terrain mapping information). Data can be transferred to DCIEM through tape, and electronically in the future.

11. **Processing.** The flight information and video data will be stored on the Aerospace Physiology Group's Epoch file server/storage unit. Video recordings will be stored in digital format and subsequently transferred to the database. This will allow real-time playback of video images on the Sparc workstations in synchrony with displays of flight data recorder information.

12. Data processing will provide a central challenge. The group plans to use the Oracle database management product already in-house at DCIEM. This database management system has the ability to handle all the ATG and Squadron data on personnel and missions, the flight information, as well as the video data. The Oracle product also has excellent security features which will allow controlled access to the stored data. The actual database specific to this study will be called HERCULES. An extensive search of the scientific literature and commercial material from software companies regarding database technologies is on-going. The literature review will identify:

- a. methods of scoring flight performance;
- a. statistical techniques for processing the data; and
- b. methods of analyses for tracking performance changes over long periods of time.

13. **Analysis.** In addition to the development of software for acquiring and storing data from both the simulator and the aircraft, a software package is under development to allow for flexible analysis of the data. The basic philosophy of our approach is to use existing software packages for file manipulation, signal processing, statistical analysis, and graphical display. All analysis and display software will run on the SPARC UNIX workstations at DCIEM. A previously developed DCIEM package, SWAP, is being re-written and re-named, to provide a flexible system for facilitating the multi-media data exploration, analysis, and presentation requirements. Through a combination of multiple graphic and text screens, pull-down and pop-up menus, icons, buttons and scroll lists, it will simplify the complex task of data analysis by providing a visual representation of the data and its processing paths.

14. A literature search is also ongoing to investigate additional tools for performing automated analysis of both flight performance and crew

performance. A number of commercial airlines worldwide collect and analyze FDR data for use in aircraft maintenance programs. In several European airlines these data are also used for the analysis of aircrew performance. Though no North American airlines have a program in place, preliminary studies have been undertaken to develop a program similar to the European approach in both Canada and the US. A meeting was held with the Transport Development Center (TDC) - Transport Canada, in Montreal, to discuss possible areas of collaboration with respect to the collection and analysis of FDR information. TDC personnel have established a formal collaboration with NASA Ames and US military operators in this area. They were able to provide the names of the key personnel involved in these types of programs in Canada, USA, and Europe.

15. The major US program is the Automated Performance Measurement System (APMS) undertaken by NASA and the FAA. The work is being performed by the Aerospace Human Factors Division of the NASA Ames Research Center. The statistical techniques for correlating crew behaviours with flight performance will be investigated along with available software for graphical displays of flight analysis, other AI techniques for analysis of flight performance, robust techniques for measuring trends, and spectral analysis. Most of these techniques would be targeted towards minimizing the personnel requirements necessary for on-going analysis and report generation.

Crew Performance Assessment Working Group

16. While this study is not restricted in the human factors issues it will consider, heavy emphasis will be placed on workload and workload management related topics. DCIEM has accumulated a substantial knowledge base in the area of human information processing and the measurement of operator workload. This knowledge base provides a framework for integrating several theoretical constructs that are relevant to the study of human information processing. This synthesis of these theories is ongoing but is already providing strong guidance to this group's activities in the development of a human performance measurement battery. The contractors are attempting to rationalize this theoretical framework with the more classical social psychology approach that has underpinned CRM/ACT programs to date.

17. The primary focus of the contractor's work has been a literature review in preparation for recommending appropriate metrics for crew performance assessment. Topics reviewed include:

- a. existing performance measures for the flight deck;
- b. group processes and dynamics, especially as they relate to decision making;
- c. crew resource management;

- d. theoretical concepts of the development and use of mental representations;
- e. knowledge elicitation techniques; and
- f. other non-invasive observation techniques and alternative tools such as voice stress analyses.

18. Two main thrusts for the development of crew performance metrics are being considered. The first involves creating verbal and non-verbal measures of aircrew performance, arising from observation of CC-130 operations, simulator work, and self-reports of aircrew. The second involves the measurement of both individual mental models of safe flight operations, and group mental model consensus.

19. Once developed, the measures will serve as indicators of safe flight performance and will be used in a variety of CC-130 studies in the simulator, and possibly in operational aircraft as well. Beyond this use, these measures will also have potential future application to DCIEM laboratory-based studies to give additional insight to what is learned from simulator experiments. This will be particularly useful for issues where high numbers of subjects are required or where ethical or operational issues might prevail. Low-fidelity simulator studies, using the measurement battery with non-ATG subjects, could help refine critical safety factors and test variations with no ethical risks or danger to participants.

20. It is also important to explore those cognitive styles that may affect decision making and group interaction on the flight deck. To this end, the contractors are reviewing this literature and propose to test several measures of cognitive style that are related to individual decision making effectiveness. Other cognitive style scales have relevance to how people might react to interpersonal situations, with implications for the group or team problem solving process as it arises on the flight deck. They will also explore the emerging literature concerning leadership and followership.

The Way Ahead

21. Data retrieval, processing, storage and analysis techniques will be developed by the Flight Performance Assessment Group. Digitally stored video data will allow a simultaneous presentation of flight animation, statistical analysis of data, and cockpit video/audio. This presentation will then be used by the Crew Performance Assessment Group to assess and score crew behaviour during simulator training tasks according to criteria under development. The consolidated method will consist of:

- a. a measure of crew performance during a simulator task; and

b. a simultaneous measure of flight path data.

22. Pilot studies on several measurement methods are planned in conjunction with the Spring 95 Box Top exercise. The purpose of these studies is the validation of metrics and the development of baseline data. The methods will include a wristwatch-sized device called the Actigraph for measuring aircrew work/rest cycles, a one page activity log, and several simple scales to measure subjective experiences of well-being, fatigue and workload. These activities will be designed to be as non-intrusive as possible, and will involve a minimum of crew time. Procedures will be cleared with ATG and individual crews before implementation.

23. Despite administrative delays in processing contracts, the 15 Oct 95 deadline for completing the method development appears to be feasible. It is recommended that the task deadline for completing the experimental design phase remain as 15 Oct 95 for the presentation of our draft report to ATG. After input is received from ATG, a final report will be issued on 15 Dec 95. The deadline for the validation experiment on Aircrew Coordination Training (ACT) will be July 96. These recommendations are reflected at Appendix 2.

24. Following the completion of the ACT study in July 96, it is recommended that additional simulator-based studies examine selected human factors issues. Previously identified ATG concerns (reference A) and other issues such as fatigue, circadian rhythm disturbance, scheduling policies, and crewing practices should be studied. As experience with the use of the technology is gained, it is anticipated that these studies will run relatively quickly and efficiently.

25. It is likely there will be a direct application of this technology to in-flight studies. There is the potential to collect data (including video) continuously during all flights. Data could be stored confidentially for selective processing at DCIEM. Though analysis of all the data that would accumulate from the operational environment may not be practical, accident trend and human factor measures could be possible.

Summary

26. The Flight Performance Assessment Working Group is developing techniques for the collection, storage, processing and analysis of flight performance (e.g., attitude, altitude, airspeed, glide slope deviation, etc.) and cockpit video data from the CC-130 Hercules Simulator. There appear to be no technical obstacles for collection and processing of flight information to a DCIEM database. An ongoing literature search will investigate tools for performing automated analysis of both flight performance and crew performance.

27. The Crew Performance Assessment Group are considering two main areas for the development of crew performance metrics. The first involves

measures of aircrew performance based on observation of CC-130 operations, simulator work, and self-reports of aircrew. The second involves the measurement of both individual mental models of safe flight operations, and group mental model consensus.

28. Digitally stored video data will allow a simultaneous presentation of flight animation, statistical analysis of data, and cockpit video/audio. This presentation will then be used to assess and score crew behaviour during simulator training tasks according to criteria under development. Once developed, the measures will serve as indicators of safe flight performance and will be used in a variety of CC-130 studies in the simulator, and possibly in operational aircraft as well.

29. The final report including ATG input will be issued on 15 Dec 95. The deadline for the validation experiment on Aircrew Coordination Training (ACT) will be July 96. Following the completion of the ACT study in July 96, it is recommended that additional simulator-based studies examine selected human factors issues.

Recommendation

30. We recommend approval of the revised Organization Plan at Appendix 1, and the revised Work Plan at Appendix 2.

Banks
LCol
Chairman, Study Group

Appendices:

Appendix 1
Appendix 2

ANNEX D

Appendix 1. Organization of the Study

(HPSD)

15 March 1995

Organization

Study Group

LCol Bob Banks - Chairman (DCIEM)

LCol John Jensen - Co-chair ATG (ATG)

LCol Mark Tysiaczny - AIRCOM Rep

Mr. Bill Noble - CRAD Rep

Maj Barry Davis - ATG Coordinator

Mr. Keith Hendy - Crew Performance Assessment

Mr. Bill Fraser - Flight Performance Assessment

Capt Helen Wright - Central Coordinator

Capt Brad Waddell - 10 TAG Observer

Crew Performance Assessment Working Group

Aim: To develop a method of precise measurement of crew performance in the CC 130.

Mr. Keith Hendy - Committee Head

Capt Brad Waddell - 10 TAG Observer

Mr. Ian Mack - Defence Scientist

Dr. David Jamieson (contractor)

Dr. Megan Thompson (contractor)

Capt Helen Wright

Other Human Factors Division/ATG staff as required

Flight Performance Assessment Working Group

Aim: To develop a method of precise measurement of crew performance in the CC 130.

Mr. Bill Fraser - Committee Head

Mr. Tom Gee - Computer Engineer

Dr. Mio Jankovic from Prior Data Sciences (contractor)

Mr. Paul Comeau - ATG Operational Researcher

Other HPSD/ATG personnel as required

ANNEX D

Appendix 2. Work Plan

Work Plan

- | | |
|---|-----------|
| 1. Evaluation of technical feasibility | 12 Sep 94 |
| 2. Hire 2 consultants on contract | 15 Oct 94 |
| 3. Literature review | Ongoing |
| 4. Familiarization flights for study group members | Ongoing |
| 5. Final design methodology | 15 Sep 95 |
| 6. Draft report on development of experimental method | 15 Oct 95 |
| 7. Final report on development of experimental method | 15 Dec 95 |
| 8. Complete an experiment on current ATG CRM | 15 May 96 |
| 9. Identify additional human factors for further eval | 15 Jul 96 |

Reporting Plans

- | | |
|--|-----------|
| 1. Initial feasibility report to SMC DCIEM and ATG | 12 Sep 94 |
| 2. Progress report to ATG/DCIEM | 15 Mar 95 |
| 3. Final report on method development | 15 Dec 96 |
| 4. Final report on CRM experiment | 15 Jul 96 |
| 5. Final report on additional human factors | 15 Jul 96 |

Start Date: 15 Oct 94

Estimated Completion Date: 15 Jul 96

ANNEX E. REPORT OF THE BOX TOP STUDY

AIRCREW WORK/REST CYCLES IN BOXTOP 1/95

ABSTRACT

This paper reports on activity log data collected during Boxtop 1/95. Boxtop aircrew flew a fixed three-shift, three-airplane schedule repeating every 48 hours. A two-page activity log was used to collect data from 41 aircrew and wrist Actigraphs were used to provide a second source of information from 18 aircrew. Information was collected on sleep length, hours awake before flying, hours continuously awake, mood, alertness and sleep quality. Significant differences were found between shifts, particularly the night shift which scored lower on mood, sleep quality, and alertness on awakening. Afternoon crews were up longer before flying and at their last touch-down of a work cycle, on average, than morning or night crews.

INTRODUCTION

Purpose. This report discusses the field data collection of work/rest cycle data from CC-130 aircrew - to validate field data collection methods and instruments; to collect baseline data on a Boxtop operation; and to provide feedback to the Commander of Air Transport Group on CC-130 human factors issues. The data were collected during Boxtop 1/95, a two-week (17-27 April, 1995) airlift operation to deliver fuel from Thule Air Force Base, Greenland, to Canadian Forces Station (CFS) Alert.

Background. On October 30, 1991 CC-130 aircraft 130322 crashed near CFS Alert during a Boxtop operation. Because of this and other accidents in the CC-130 fleet, a human factors study was sponsored by the Canadian Forces (CF) Air Command (AIRCOM). The study is the joint responsibility of Air Transport Group (ATG) and the Defence and Civil Institute of Environmental Medicine (DCIEM). The study mandate centred on developing measures of aircraft and aircrew performance to be tested and validated using the CF CC-130 simulator. The effort reported herein is a collateral report to the main study in response to an early request from ATG that DCIEM provide informed comment on their guidelines for work/rest cycles. It is recognized that other important human factors issues are of concern to ATG personnel.

The Canadian Forces Air Transport Group dedicates three CC-130 Hercules aircraft to each Boxtop operation, along with several crews. For Boxtop 1/95, three shifts of three crews were scheduled and within a shift, crews were called out one and one-half hours apart to reduce the possibility of congestion on the ramp in Alert. The 'morning' shift call-out times were 6:00 a.m., 7:30 a.m. and 9:00 a.m., followed by the 'night' shift call-outs at 10:00 p.m., 11:30 p.m., and 1:00 a.m. The final three crews in the 48-hour cycle - the 'afternoon' shift - were

called at 2:00 p.m., 3:30 p.m., and 5:00 p.m. the day after the morning shift began (see Appendix 1 for the Boxtop 1/95 crew call-out schedule).

Ideally, each crew flies three round-trips during a shift. Aircraft shuttle around the clock on the 90-minute flight between Thule and Alert. Weather and breakdowns often affect the flow of the operation and the crew duty days. In Boxtop 1/95, 27 trips were canceled because of bad weather and 7 were canceled because of aircraft problems, out of a planned 145 trips².

Boxtop aircrew are drawn from the six squadrons that fly the CC-130 aircraft in different roles (strategic and tactical transport, rescue, and training). The Boxtop operations staff, which works out of Thule, also makes up a spare aircrew in the event that one person cannot fly; during Boxtop 1/95 crew substitutions occurred twice.

Aircrew work/rest cycles in long-haul operations have been of concern for many years [1,2,3]. The condition known as fatigue is a typical result of poor work/rest schedules. To paraphrase Bartlett [4], fatigue can be defined as:

a deterioration in performance over time, under normal conditions, that leads to unwanted results.

Fatigue has been classified into several types (e.g. [1,2]), however it is generally accepted that mental fatigue and physical fatigue have different causes and results. In flying, mental fatigue is the greater threat to safe operations³ and as described in [1] the results can include: a decrease in the ability to recognize a changing situation; delays in correcting a situation once it has been recognized; and increased 'sloppiness' in making corrections.

More recently, Belland and Bissell [5] examined fatigue in U.S. Navy flying operations during post Gulf War patrols of the southern no-fly zone over Iraq. Survey data were collected on fatigue for 125 aircrew that flew four to six hour sorties daily, in single and dual seat jet aircraft (e.g. F/A-18, A-6.) Several physical symptoms of fatigue were reported by aircrew, including headache, back pain, and drowsiness. Aircrew also reported increases in small mistakes and insomnia. Aircrew reported that having a 'no fly' day every four or five days helped them catch up lost sleep.

The measurement of aircrew performance in-flight, without interfering with the task at hand, presents significant difficulties. In one early study by McFarland and Edwards [6], the authors took advantage of extra crew to administer physiological and psychological tests on a trans-Pacific flight. However, in

²In this paper, 'trip' and 'round-trip' are used interchangeably.

³Bartlett hypothesized that mental fatigue effects become significant long before physical symptoms occur.

modern military flying, there are few opportunities available to perform such comprehensive testing in the field. More recent studies have concentrated on the use of questionnaires and tests administered on the ground.

DCIEM also has experience studying work/rest cycles in long range airlift operations. A 1970 study by Innes [3] used a fatigue checklist to make pre- vs. post-flight comparison for several long flights (legs of 7 to 12 hours duration) in Hercules and Yukon aircraft. Results showed that subjective fatigue information could be distinguished using a questionnaire format. A more recent Actigraph study of long range CC-130 operations was performed by Donati[7]. These, as yet unpublished, data provided experience with the use of Actigraph monitors for studying aircrew in an operational environment. In the Donati study, a single aircrew was monitored during long haul flying between Canada and Somalia.

The invasion of Kuwait in the summer of 1990 led to a massive military build-up in Saudi Arabia, including a significant number of long range flights by military aircraft. The buildup and subsequent conflict provided ample opportunity for field studies of aircrew work/rest cycles.

Neville et al. [8] reported on subjective fatigue of USAF C-141 crews during the Persian Gulf war. Aircrews flew nominal 16 hour days (unaugmented) which were often extended to 20 hours. Flight records, activity logs, temperature data, and aircraft digital flight data were collected; and two fatigue scales were administered. The authors found that fatigue was related to 48-hour flight history; 10 hours of sleep in a 48 hour period was not sufficient for recovery. More than 15 hours of flying in a 48 hour period was also linked to high fatigue. The authors identified the need for careful 'fatigue management' in airlift operations, including (1) paying close attention to on-the-ground and on-board sleeping facilities for aircrew, and (2) fatigue management training for transport aircrew.

In a similar study of USAF C-5 operations during the force build up for the Persian Gulf War (Operation Desert Shield), Bisson et al. [9] found comparable results. Activity logs and flight records were used to collect information on crew performance. Aircrew reported moderate to extreme fatigue ratings on cross-ocean and round-trip flights. Periods of 8 to 18 hours cumulative airborne time were reported in a 24-hour period. Aircrew with late night take-off times reported being awake an average of 10.3 hours before flying. Difficulties falling asleep and restless sleep were also common complaints of aircrew required to make significant circadian shifts.

These previous studies of aircrew work/rest cycles in operational settings provide a wealth of suggestions for test instruments [5,8,9]. Two instruments were selected to measure work/rest data: Neville's questionnaire and activity monitors (Actigraphs).

METHOD

Subjects. Boxtop 1/95 assigned aircrew consisted of 45 adults - 43 males and 2 females. Participation in this study was voluntary and data were received from 41 people (91% participation). CC-130 aircrew positions include three officers: the Aircraft Commander (AC); the Co-Pilot (CP); and the Navigator (NAV). There are two other ranks positions to complete the standard CC-130 crew of five: the Flight Engineer (FE) and the Loadmaster (LM). Four crew members have cockpit stations; the Loadmaster works primarily in the rear of the aircraft.

Aircrew ranks ranged from Master Corporal (MCpl) through to Major (Maj) and years of military service⁴ ranged from 3.5 to 31. Experience differences exist among the crew members: Flight Engineers on the CC-130 are required to have experience on other aircraft before converting to the Hercules, and they must have attained the rank of Sergeant. Under recent changes to ATG regulations, pilots cannot become a CC-130 Aircraft Commander during their first tour (a 'tour' or 'posting' is typically 3 to 5 years). The Navigator, Loadmaster, and Co-Pilot Positions may all be filled by personnel who have just completed basic training (in the case of the CP position, this means basic flying training, plus the CC-130 conversion course.)

Total military flying experience reported by Boxtop 1/95 aircrew ranged from 450 hours to 8400 hours, with a mean of approximately 2750 hours and a standard deviation of approximately 2025 hours. Reported CC-130 (type) experience ranged from 100 to 7500 hours, with a mean of approximately 1700 hours and a standard deviation of approximately 1630 hours. On average, the AC was the most experienced member of the crew on type, although there was no significant difference between AC experience and LM experience in pairwise comparisons. The CP was the most type-inexperienced member of the CC-130 crew. Mean flying hours on the CC-130 are presented in Table E1.

Table E1. Mean CC-130 Experience for Boxtop 1/95 Aircrew (hours) — sample size in parentheses

Aircraft Commander (8)	2550
Co-Pilot (7)	281
Navigator (8)	885
Flight Engineer (9)	2200
Loadmaster (9)	2286

Boxtop 1/95 crews were scheduled to work a nominal 16 hour day⁵, followed by 32 hours of rest. Crew rest is defined as any time not spent working, and

⁴'Tombstone' data on air crew were collected from two questionnaires handed out at Boxtop, but not discussed in this paper.

⁵ Under ATG regulations, the maximum allowable crew day is 18 hours.

includes both sleep and wake time. Crews were scheduled to start in three groups of three, as shown in Table E2. For each group, crew call-out times were staggered to reduce the chance of congestion on the ground at CFS Alert. Using nine crews and three aircraft results in a 48 hour cycle.

Procedure. Activity monitors were given to four of nine crews participating in Boxtop 1/95. Of 20 possible subjects, 18 volunteered to wear activity monitors (model AMA-32, by Precision Control Design Inc., Fort Walton Beach, Florida.) Additionally, aircrew were asked to complete a daily activity log, described below.

Subjects were allowed to wear the activity monitors on either wrist, using a light nylon watch strap. They were asked to wear the Actigraph at all times, including sleep periods, except for showering and unusually vigorous exercise. Subjects were asked to note on their activity logs periods when the activity monitors were not worn.

Table E2. Boxtop 1/95 Crew Call-Out Schedule

Crew	Call-Out Time
51	6:00 a.m., Day 1 (morning shift)
61	7:30 a.m., Day 1 (morning shift)
92	9:00 a.m., Day 1 (morning shift)
41	2:00 p.m., Day 2 (afternoon shift)
31	3:30 p.m., Day 2 (afternoon shift)
91	5:00 p.m., Day 2 (afternoon shift)
21	10:00 p.m., Day 1 (night shift)
62	11:30 p.m., Day 1 (night shift)
52	1:00 a.m., Day 2 (night shift)

The second instrument used to collect data on work/rest cycles was an activity log. A two-page activity log used by Neville et al. [8] was adapted for the study and reproduced using two sides of a single sheet of paper for each day of the operation. An example is given at Appendix 2. In order to encourage continuing participation by aircrew, it was decided to give out the logs at the beginning of every crew shift (i.e. every two days.) Initially this was done by handing ten blank copies to the Aircraft Commander as he reported to Boxtop operations at the start of a new shift. After the aircrew were familiar with the sleep logs, they were inserted into the satchel given to each Aircraft Commander at the beginning of his shift. Additional logs were kept in the operations area for use as needed.

All 45 aircrew had the opportunity to participate in the activity logs. Actual participation was 41. Participants were asked to identify themselves using the last three digits of their service number plus their initials. Identification was necessary to compare the three different crew start times. Some individuals were concerned that use of their service numbers would compromise their

anonymity and as a result only 32 (71%) of the aircrew who participated could be assigned to a crew.

DIFFICULTIES ENCOUNTERED WITH FIELD DATA COLLECTION

Actigraphs. Because of the duration of the operation, it was decided to perform a data down-load from the Actigraphs at approximately half-way through the Boxtop operation. The devices were retrieved from the aircrew, down-loaded, and the batteries replaced. The Actigraphs were then re-initialized and returned to the aircrew. Since four crews were involved, representing each of the three groups of scheduled starting times, the Actigraphs could not be returned immediately. Due to difficulties down-loading data from the Actigraphs, data were obtained from only one of the 18 devices at the half-way point.

The Actigraphs were checked again after four days. No data could be down-loaded at that point, so the devices were not returned to the aircrew. After returning to DCIEM, it was discovered that the Actigraphs had functioned properly during the four day period after the first down-load was attempted. It appears that difficulties with the computer used for the down-load were the source of the unsuccessful data transfer. The Actigraphs had been sent out for service four months before the field trial and were not returned to DCIEM until the week before Boxtop. Accordingly little time was available to test the updated equipment.

Activity Logs. Problems with the format of the sleep log became apparent quickly. The long crew days and 48 hour crew cycle led to confusion about how to record data. Some subjects took more than one significant sleep period in 24 hours. Others did not know which meals to identify as breakfast, lunch, and supper. For example, aircrew starting at 2:00 p.m. might have their first meal of the day at 1:00 p.m. - is it breakfast or lunch? For many subjects, their work day spanned two calendar days. Deciding which date to enter on each activity log was also unclear. When these problems were identified to the experimenters, subjects were told to try and be consistent and make notes on the logs to explain their responses. For most participants, the result was usable, albeit difficult to interpret information. However, with night crews (those with call out times at 10:00 p.m., 11:30 p.m., and 1:00 a.m.) it was difficult to determine whether the day they were reporting on was a flying day or a rest day.

The activity logs were often not completed first thing in the morning or last thing at night as requested at the top of the page. Some crews completed the forms in the air between Thule and Alert. Most of the requested information was provided, but some was omitted and some was entered inconsistently. For example, some subjects used 24-hour times; others used 12-hour times; still others alternated between the two systems. Data inconsistencies were found: for example, the 'wake' page asked about the number of times a subject awakened during the night. Subsequent questions asked aircrew to say 'how many times

for bathroom', 'how many times for noise', etc. The sum of these subsequent questions did not always equal the response to the first question. Subjects' sleep periods were identified by asking: time in bed, lights out time, minutes until fell asleep, and wake-up time. In many instances, at least one of these was omitted, reducing the number of valid responses.

Environment. Finding an opportunity to brief all the aircrew was difficult. The study was introduced and briefly explained as part of the operational briefing the day before Boxtop began. However, the available time was restricted and a well organized and speedy briefing was necessary. On a Boxtop operation, accommodation in either Thule or Alert is difficult to obtain, requiring a minimum of personnel. Arranging for someone to be present to collect sleep logs at each crew call-out was difficult, and some pick-ups were missed.

RESULTS AND DISCUSSION

Actigraphy. Eighteen aircrew wore Actigraphs and data were retrieved from 15 of 18 for a four day period during Boxtop. Sleep periods were determined from the Actigraph data using a standard algorithm [10]. The resulting graphs were compared with activity log data reported by the aircrew. On a case by case basis, subjects appeared to be good at estimating their lights out time (i.e. the time they began trying to go to sleep) and their wake-up time - differences of up to 15 minutes were noted. However, they appeared to be less able to estimate the time it took to fall asleep, used in the calculation of estimated sleep length. Significant discrepancies were also seen in the number of times awake, and the amount of time awake. However, it cannot be established whether these differences were due to subjects' estimation abilities or uncertainty in the Actigraph sleep scoring algorithm.

The Actigraph count data⁶, recorded at 30 second intervals, were selected for incidences of three-trip flying days. For each crew, data were available for exactly one three-trip day. Analysis of variance was used to compare mean activity levels within shifts for the three trips; individual and shift analyses were performed. Table E3 presents the results. All three shifts showed significant differences⁷ in activity levels among trips.

Table E3 shows changes in activity levels from trip to trip with each crew. Morning and night crew activity levels drop from first to second trip, then climb again for the third; the third trip activity level does not reach the first trip level. In contrast, afternoon aircrew decline in activity level from first to second trip,

⁶Actigraphs make 0/1 assignments of movement at a fixed sampling rate, and sum over a pre-determined interval (in this case, 30 seconds) to produce a 'count' that is recorded.

⁷All differences between data sets were tested using the SAS system analysis of variance test. 'Significant differences' are those which are due to chance less than 5 times in 100.

and from second to third trip. Activity level trends of morning and night crews support comments reported by Boxtop 1/95 aircrew that the second of three trips was the 'hardest'. No significance should be drawn from activity level comparisons among shifts for a given trip. While the activity values show large variation, this could be the result of variation among the Actigraphs, and/or individual differences (small sample size).

Table E3. Mean Activity Scores by Shift

Shift	Trip 1	Trip 2	Trip 3
Morning	58.00	56.71	57.59
Afternoon	50.02	49.77	47.39
Night	53.04	50.53	52.64

Since Actigraphs measure movement, background vibration may affect the performance of the device. If background effects are too great, it may not be possible to determine, for example, periods of sleep in an aircraft. In a recent paper [13], Sadeh et al. noted that externally induced movement can affect Actigraph recordings to the point where a person sleeping in a moving vehicle cannot be distinguished from one who is awake. While this latter point could not be confirmed, significant background effects were observed. On the return trip from Thule, all the Actigraphs were packed into a padded case which was shipped by CC-130 and each printout shows a significant trace for all airborne periods (see Appendix 3: the period from 0800 hours on the day identified by trace line number five on the y-axis of the graph through 0130 hours on the day identified by trace line 6 on the y-axis of the graph). The Cole and Kripke sleep scoring algorithm could not identify the devices as being stationary. However, this factor is probably not significant for Boxtop 1/95 data due to the short length of each flight.

Sleep Cycles. Aircrew did not go to bed at the same time every night; nor did they sleep for the same number of hours each night. From the activity logs collected, 319 sleeps could be identified over all subjects and all days of the operation. Estimated sleep time⁸ ranged from 30 minutes to 13 hours and 50 minutes for the longest single sleep period in 24 hours. The mean value was 7 hours 11 minutes, with a standard deviation of 2 hours 16 minutes.

When the estimated sleep data are divided up by aircrew shift, the number of available records drops to 261. The mean sleep length by shift and standard deviation are shown in Table E4. Mean estimated sleep lengths were compared using analysis of variance and found to be significantly different.

⁸Sleep time was estimated by wake-up time minus lights out time minus time to fall asleep minus total time awake during the sleep. Actigraph comparisons indicate that subjects were good at estimating lights out time and wake-up time, but varied in their estimates of time to fall asleep.

Table E4. Mean Estimated Sleep Length by Shift

Shift	Mean	Std Dev	Sample Size
Morning	7 h 35 min	1 h 48 min	103
Afternoon	7 h 10 min	1 h 59 min	81
Night	6 h 25 min	3 h 32 min	77

Mean sleep length for the first five days (17-21 April) was not significantly different from that of the last five days (23-27 April) of Boxtop 1/95 in total, or by shift⁹. However, examination of individual data shows sleep lengths alternating between flying and resting days. Figure E1 shows sleep length vs. date for Subject 001. Long sleeps on even numbered days occurred after completion of a flying cycle.

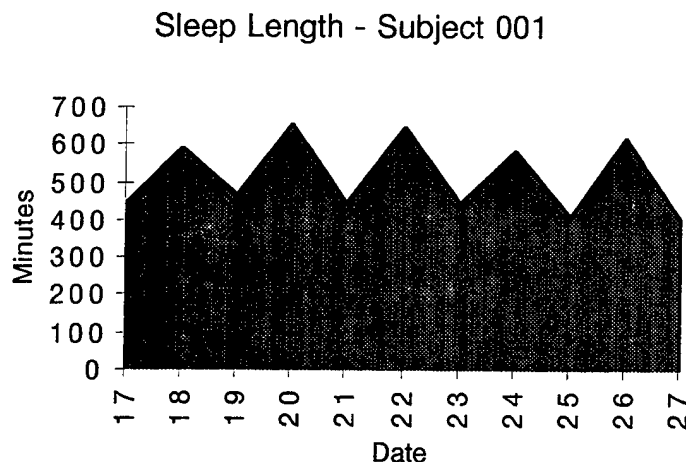


Figure E1. Sleep Length vs. Date for Subject 001

Each subject's estimated sleep lengths were divided into two groups, one for sleep after flying and one for sleep after a rest day. Some data were omitted because it was not possible to determine which day was flying vs. rest; others were deleted because of duplicate dates. A total of 252 valid sleeps remained. Mean estimated sleep after flying was 8 hours 20 minutes, with a standard deviation of 1 hour and 2 minutes. Mean estimated sleep after a rest day was 6 hours 17 minutes with a standard deviation of 1 hour and 49 minutes. The two means are significantly different.

Significant variation in sleep patterns has been recognized as contributing to decreased performance in flying [11]. As the amount of sleep decreases, flight crew can expect to take longer to recognize warning cues, understand their

⁹April 22 was a flying day. It is omitted here to balance the number of days compared.

significance and act on them. As noted earlier, Neville et al. [8] found that fatigue was related to 48-hour history of sleep and flight time. Ten hours sleep or less in 48 hours was not sufficient protection against fatigue, and 15 hours or more flight time was associated with high subjective fatigue ratings. On average, Boxtop aircrew received sufficient sleep, but the long crew days experienced in some cases are a warning sign of potential fatigue.

Time awake before first take-off; Time continuously awake. A long time awake before the start of the duty day is an indication that aircrew are not adjusting to an unusual shift. Time awake before duty was estimated by the difference between wake time and time of first take-off on a flying day. Using first take-off time is a relatively poor estimator of the beginning of the duty day. A better estimator would be the crew call-out time; however these were not recorded. Under routine transport operations the aircrew, especially the flight engineer, may be on duty several hours before the take-off time. However, during Boxtop operations, the aircraft are usually not available for pre-flight activity too early, and operations staff call out crews with the intention of a minimum wait before first take-off.

From the sleep log data, hours awake before take-off could be calculated in 115 instances. These data were analyzed using analysis of variance and grouping by shift. Mean time awake before first take-off is shown in Table E5. The means are significantly different. Aircrew on the afternoon shift were up the longest before flying: an average of 5 hours 4 minutes. It was initially expected that night crews would be up longer before flying, because of disturbed circadian rhythms. On closer examination of the data, it was found that, in contrast to their colleagues, night crews often napped for two to four hours before flying.

Table E5. Mean Time Awake Before First Take-Off

Shift	Mean	Std Dev	Sample Size
Morning	2 h 47 min	1 h 17 min	46
Afternoon	5 h 4 min	2 h 55 min	44
Night	3 h 16 min	1 h 47 min	25

Using the same set of data, it is possible to estimate the time continuously awake when aircrew make their last landing. As an estimator for this value, the last landing time from the K-1017 form¹⁰ was added to the time awake before flying, less thirty minutes to approximate the time to begin final approach. Data were analyzed by number of trips flown. Results for three trip days are presented by shift in Table E6.

¹⁰The K-1017 form is Air Transport Group's 'Flight Authorization and Record of Flight' form. It is completed for each flight of a CC-130 aircraft. More detail on the K-1017 is available from the author.

Table E6. Estimated Mean Time Awake on Last Final Approach (3 trip days)

Shift	Mean	Sample Size
Morning	15 h 4 min	24
Afternoon	17 h 19 min	35
Night	16 h 11 min	10

While mean values show typical aircrew condition, it is interesting to look at the extreme cases for hours awake on last landing. Using touch-down times from the K-1017 form and reported wake-up times for aircrew, the four largest values for continuous hours awake on last touch-down are: 23 hours 55 minutes, 24 hours 35 minutes (2 cases), and 25 hours 5 minutes. The ten largest values of continuous time awake are all from crews on the afternoon shift. Note that afternoon crews flew, on average, more trips per day than did morning or night crews - because of weather and aircraft maintenance problems - so there was more opportunity for these crews to have been up for long hours on their last landing.

Sleep Quality. While the mechanisms are not well understood, mood, sleep quality, and alertness on awakening are important indicators of fatigue (e.g. [8, 12]). On their activity logs, aircrew were asked to rate sleep quality, mood on awakening, and alertness on awakening on a seven point scale. Adjectives were provided for the lowest and highest values (see Appendix 2) and all scales associated a score of one (1) with the poorest rating and seven (7) with the best rating.

Comparative data were not collected before or after Boxtop, so the absolute ratings cannot be commented on. However, changes in ratings and differences between shifts are of interest. Table E7 presents results on a per shift basis. Analysis of variance established statistically significant differences among the means on all three scales. While pairwise testing was not done, the significances seem to be due to the night shift which scored lower, on average, on all three scales. Night crews were not sleeping as well, were not as positive, and did not feel as alert as morning and afternoon crews.

Table E7. Mean Mood, Sleep Quality, and Alertness Ratings by Shift

Shift	Mood	Sleep Quality	Alertness on Wakening
Morning	5.05	4.84	4.77
Afternoon	5.08	4.81	4.73
Night	4.02	4.03	3.91

Mood, sleep quality and alertness on awakening were also compared for the first

few days (17-21 April) vs. the last few days (23-27 April) of Boxtop. No significant differences were found, however mean values for the night shift aircrew were lower on all three scales for week two as compared to week one. In contrast, shift two aircrew showed improvement on all three scales between week one and week two.

Another, somewhat more objective measure of sleep quality is number of times a subject awakened during sleep. This measure has been associated with circadian shifts and rapid shifts in sleep schedules [9]. Boxtop aircrew were asked to report number of times awake and the reasons. Of 370 valid sleeps, aircrew reported one or more wake episodes in 262 or 81% of cases. Fully 25% of sleeps were interrupted three or more times. Subjects who woke up one or more times were awake, on average, for a total of 31 minutes. Subjects were asked to say why they woke up: either for the bathroom, noise, discomfort, or other reasons. For reported values 29% were due to noise or discomfort. It was observed (and experienced!) by the experimenters that many complaints were due to the quality of accommodation at Thule Air Force Base.

COMMENTS AND CONCLUSIONS

Although problems were encountered in data collection, this study succeeded in showing that activity data can be collected in the field. This success was due in no small measure to the excellent cooperation received from the Boxtop operations staff. Data collection instruments must be tested and customized before taking them to the field to avoid unforeseen difficulties in equipment usage and questionnaire format. The benefits of a field study are operational relevance, increased impact on decision makers, and a stronger commitment from aircrew.

In engineering, fatigue is defined as failure resulting from repeated applications of load [14]; it is also called *progressive* failure. In Boxtop operations, aircrew experience long working days over a period of up to two weeks. For aircrews, repeatedly approaching their crew day limit over an extended period of time provides a good environment for fatigue 'failures' to occur.

Baseline data were collected on an ATG operation. Evidence was found (some of it compelling) that in Boxtop 1/95, afternoon crews and night crews were disadvantaged, as compared to morning crews. While night crews tended to nap before flying, they reported significantly lower mood, sleep quality and alertness on awakening - indicating that the napping strategy was not completely effective. Afternoon crews did not nap, flew more trips, and possibly as a result, were more likely to have been awake longer than morning or night crews.

The data analysis in this report has focused primarily on mean values of measures such as sleep length and hours awake before flying. However, on average, every flight is expected to be incident free. The analysis of data for flight

safety must include consideration of the extreme cases - because accidents are almost always the result of conditions that statistically would be considered as 'outliers'. This was illustrated by highlighting the very long days experienced by some crews.

Flying contains an element of risk, and the flight safety system exists to manage this risk. Using an analogy from the field of engineering, there is a safety margin between the operational needs of a mission and an incident or accident. Maintaining this margin is the responsibility of all members of the aviation community, including researchers. Prolonged flying operations should be monitored for symptoms of crew fatigue and a narrowing of the margin of safety. This study has investigated one way to add to the knowledge base of the CF aviation community and it has provided baseline data for future comparison.

This study, on its own, does not provide sufficient evidence of aircrew fatigue, or other flight safety considerations, to justify a reduction in the length of the Boxtop aircrew day. However, the data do suggest that the three trip day may be too long. A study comparing data from a two-trip-day Boxtop and typical ATG squadron operations would provide better evidence for decision makers.

ACKNOWLEDGMENTS

The author would like to acknowledge the significant contribution of Mr. Keith Hendy in planning the field trial and providing advice on the experiment. The assistance of M. Pierre Fournier, Ms. Patti Odell, Ms. Yvonne Shek and Dr. Megan Thompson was appreciated in collecting the activity logs. Patti Odell is further recognized for the suggestion to obtain copies of the Boxtop flight records (K-1017 forms) which provided valuable information for the analysis. The support of the Operational Research Advisor of Air Transport Group, including M. Fournier's participation, is appreciated. Dr. Jon French of the USAF's Armstrong Laboratories provided the activity log used in the study.

The excellent cooperation provided by Maj Barry Davis of ATG and the Boxtop operations staff led by Maj Tom Whitburn was essential to the success of this field trial. Without out the help of the ops staff - including Capt Andy Tissot, Capt Bob Copeland, Capt Stephane Isabelle, 2Lt Shane Roberts, Sgt Al Allaire, and Sgt Bob Andrews - this study would never have happened. And finally, a special thanks to the Boxtop aircrew who took the time to participate in this study, despite the long work days.

REFERENCES

1. McFarland, R.A., Human Factors in Air Transportation, New York: McGraw-Hill 1953.
2. Innes, Capt L.G., Aspects of Fatigue in Aviation, Canadian Forces Institute of Aviation Medicine, Report 68-RD-6, 1968.
3. Innes, Capt L.G., Nervous Fatigue in Long Range Aircraft Crews, Canadian Forces Institute of Environmental Medicine, Report 70-RD-2, 1970.
4. Bartlett, F., Psychological Criteria of Fatigue, in Symposium on Fatigue, Floyd, W.F., and Welford, A.T., The Ergonomics Research Society, London: H.K. Lewis, 1953.
5. Belland, K.M., and Bissell, C., A Subjective Study of Fatigue During Navy Flight Operations Over Southern Iraq: Operation Southern Watch.
6. McFarland, R.A. and Edwards, H.T., The Effects of Prolonged Exposures to Altitudes of 8,000 to 12,000 feet during Trans-Pacific Flights, Journal of Aviation Medicine, 8(4):3-24, 1937.
7. Donati, Capt L., Personal Communication, 15 June 1995.
8. Neville, K.J., Bisson, R.U., French, J., Boll, P. and Storm, W.F., Subjective Fatigue of C-141 Aircrews during Operation Desert Storm, Human Factors, 1994, 36(2), pp.339-349.
9. Bisson, R.U, Lyons, T.J., and Hatsel, C., Aircrew Fatigue During Desert Shield C-5 Transport Operations, Aviation, Space, and Environmental Medicine, Sept 1993, pp.848-853.
10. Elsmore, Tim, General Activity Program, Office of Military Performance Assessment Technology, Washington, DC.
11. Flight International, Fatigue in the context of flight safety, 2 September 1971, pp.362-365.
12. Totterdell, P., Reynolds, S., Parkinson, B., and Briner R.B., Associations of Sleep With Everyday Mood, Minor Symptoms and Social Interaction Experience, Sleep, 17(5), pp.466-475, 1994.
13. Sadeh, A., Sharkey, K.M., and Carskadon, M., Activity-Based Sleep-Wake Identification: An Empirical Test of Methodological Issues.

14. DeGarmo, E.P., Materials and Processes in Manufacturing, Macmillan Publishing, New York: 1979.

ANNEX E

Appendix 1. Boxtop 1/95 Crew Callout Schedule

----- Morning Crews ----- ----- Night Crews ----- ----- Afternoon Crews -----

51	61	92	21	62	52	41	31	91
17/0600	17/0730	17/0900	17/2200	17/2330	18/0100	18/1400	18/1530	18/1700
19/0600	19/0730	19/0900	19/2200	19/2330	20/0100	20/1400	20/1530	20/1700
21/0600	21/0730	21/0900	21/2200	21/2330	22/0100	22/1400	22/1530	22/1700
23/0600	23/0730	23/0900	23/2200	23/2330	24/0100	24/1400	24/1530	24/1700
25/0600	25/0730	25/0900	25/2200	25/2330	26/0100	26/1400	26/1530	26/1700
27/0600	27/0730	27/0900						

Notes:

1. Times are planned callouts; actual times varied due to delays.
2. Each callout represents the beginning of a planned 16-hour duty day, including three Thule-Alert return trips. Not all trips were flown.
3. Top line of table is the crew number; crew numbers are derived from squadron numbers, e.g. crew 92 is the second crew from 429 squadron.
4. The table uses military style date-time groups, e.g. 17/0600 means the planned callout time for this crew is 6 a.m., 17 April, 1995; 26/1700 means 5:00 p.m., 26 April, 1995.

ANNEX E

Appendix 2. Activity Log

SLEEP DIARY : BEDTIME KEEP BY BED

PLEASE FILL OUT THIS PAGE LAST THING AT NIGHT

Day _____ Date _____ ID code _____

Today, what time did you have: Breakfast _____ Lunch _____ Dinner _____

How many of the following did you have in each time period? (if none, write '0')

	before or with breakfast	after breakfast before/with lunch	after lunch before/with dinner	after dinner
caffeinated drinks	_____	_____	_____	_____
alcoholic drinks	_____	_____	_____	_____
cigarettes	_____	_____	_____	_____
cigars/pipes/plugs (of chewing tobacco)	_____	_____	_____	_____

Which drugs and medications did you take today? (prescribed & over the counter)

Name	Time	Dose
_____	_____	_____
_____	_____	_____
_____	_____	_____

What exercise did you take today? (if none check here _____)

start	end	type
_____	_____	_____
_____	_____	_____

How many daytime naps did you take today? (if none, check here _____)

give times for each:	start	end
_____	_____	_____
_____	_____	_____

RATINGS: (please circle the number that best reflects your present state)

MOOD AT BEDTIME:

1	2	3	4	5	6	7
very negative						very positive

ALERTNESS AT BEDTIME:

1	2	3	4	5	6	7
very negative						very positive

Why do you believe that you feel this way in terms of (please be brief):

YOUR MOOD:

YOUR ALERTNESS:

SLEEP DIARY : WAKE TIME KEEP BY BED

PLEASE FILL OUT THIS PAGE FIRST THING IN THE MORNING

Day _____ Date _____ ID code _____

went to bed last night at: _____ lights out at: _____ minutes until fell asleep: _____ finally woke at: _____

awakened by (check one) alarm clock/radio: _____ call out _____ noises _____ just awoke _____

after falling asleep, woke up this many times during the night (circle):
0 1 2 3 4 5 or more

total number of minutes awake: _____

- woke to use bathroom (circle # times)

0 1 2 3 4 5 or more

- awakened by noises/other people (circle # times)

0 1 2 3 4 5 or more

- awakened due to discomfort of physical complaint (circle # times)

0 1 2 3 4 5 or more

- just woke (circle # times)

0 1 2 3 4 5 or more

RATINGS (please circle the number that best reflects your present state)

SLEEP QUALITY

1 2 3 4 5 6 7
very bad very good

MOOD ON FINAL WAKENING:

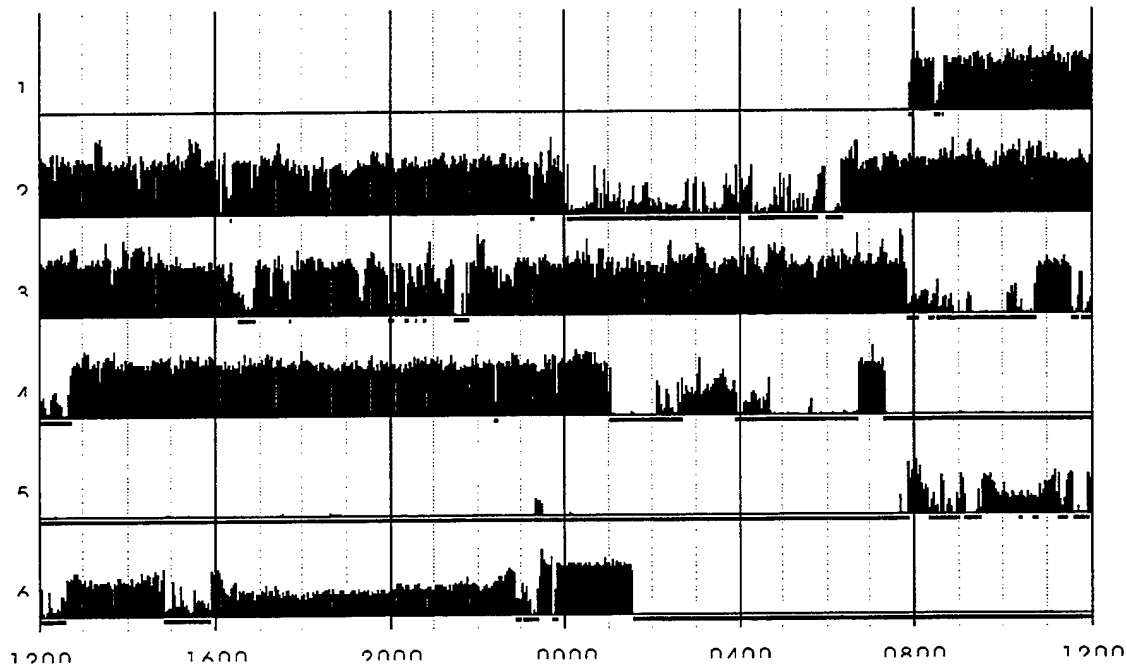
1 2 3 4 5 6 7
very negative very positive

ALERTNESS ON FINAL WAKENING:

1 2 3 4 5 6 7
very sleepy very alert

ANNEX E

Appendix 3. Sample Actigraph Plot



File: A30.DAT Epoch: 30 Amplifier: 18 Scale: 110
Starting Time: Apr 22 1995 07:55:00 SN:0030
Sleep scoring algorithm: Cole and Kripke

Explanation: An actigraph, or activity monitor, records an activity 'count' over a pre-set time interval (30 seconds in this example.) The higher the activity over the interval, the larger the count. This graph is a pictorial representation of the basic count data from an activity monitor. The trace on this graph begins at the time the activity monitor was programmed to start recording data (April 22, 1995, 07:55). On this graph, the trace continues until the memory of the device was filled, about four and one-half days later.

This graph also shows the results of "sleep scoring" using the Cole and Kripke algorithm: portions of the trace underscored with a heavy line indicate periods of scored sleep. Periods when the activity monitor was not worn are identifiable by a flat trace (e.g. Day 5, 1200 - 1600.) A restless sleep is indicated when the sleep score line is broken in several places (e.g. Day 2, 0400 - 0600.)

UNCLASSIFIED
SECURITY CLASSIFICATION OF FORM
(Highest classification of Title, Abstract, Keywords)

DOCUMENT CONTROL DATA

(Security classification of title, body of abstract and indexing annotation must be entered when the overall document is classified)

1. ORIGINATOR (the name and address of the organization preparing the document. Organizations for whom the document was prepared, e.g., Establishment sponsoring a contractor's report, or tasking agency, are entered in section 12.) Defence and Civil Institute of Environmental Medicine		2. DOCUMENT SECURITY CLASSIFICATION (overall security classification of the document including special warning terms if applicable) UNCLASSIFIED												
3. DOCUMENT TITLE (the complete document title as indicated on the title page. Its classification should be indicated by the appropriate abbreviation (S,C,R or U) in parentheses after the title.) Defence and Civil Institute of Environmental Medicine and Air Transport Group Human Factors Study of CC-130 Operations														
4. DESCRIPTIVE NOTES (the category of the document, e.g., technical report, technical note or memorandum. If appropriate, enter the type of report, e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.) Final Report														
5. AUTHOR(S) (Last name, first name, middle initial. If military, show rank, e.g. Burns, Maj. Frank E.) Banks, LCol Robert D., Hendy, Keith C., Fraser, William D., Thompson, Megan M., Jamieson, David, Wright, Captain Helen L., Gee, Tom, Mack, Ian M., Latulippe, Maj John, Davis, Barry and Cole, Mike														
6. DOCUMENT DATE (month and year of publication of document) Nov 1996	7.a. NO. OF PAGES (total containing information. Include Annexes, Appendices, etc.) 79	7.b. NO. OF REFS. (total cited in document) 6												
8.a. PROJECT OR GRANT NO. (if appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant) 03F01-11	8.b. CONTRACT NO. (if appropriate, the applicable number under which the document was written)													
9.a. ORIGINATOR'S DOCUMENT NUMBER (the official document number by which the document is identified by the originating activity. This number must be unique to this document.) DCIEM No. 96-R-66	9.b. OTHER DOCUMENT NO.(S) (any other numbers which may be assigned this document either by the originator or by the sponsor.)													
10. DOCUMENT AVAILABILITY (any limitation on further dissemination of the document, other than those imposed by security classification) <table style="width: 100%; border: none;"><tr><td style="width: 30px; border: 1px solid black;"><input checked="" type="checkbox"/></td><td>Unlimited distribution</td></tr><tr><td style="border: 1px solid black;"><input type="checkbox"/></td><td>Distribution limited to defence departments and defence contractors; further distribution only as approved</td></tr><tr><td style="border: 1px solid black;"><input type="checkbox"/></td><td>Distribution limited to defence departments and Canadian defence contractors; further distribution only as approved</td></tr><tr><td style="border: 1px solid black;"><input type="checkbox"/></td><td>Distribution limited to government departments and agencies; further distribution only as approved</td></tr><tr><td style="border: 1px solid black;"><input type="checkbox"/></td><td>Distribution limited to defence departments; further distribution only as approved</td></tr><tr><td style="border: 1px solid black;"><input type="checkbox"/></td><td>Other</td></tr></table>			<input checked="" type="checkbox"/>	Unlimited distribution	<input type="checkbox"/>	Distribution limited to defence departments and defence contractors; further distribution only as approved	<input type="checkbox"/>	Distribution limited to defence departments and Canadian defence contractors; further distribution only as approved	<input type="checkbox"/>	Distribution limited to government departments and agencies; further distribution only as approved	<input type="checkbox"/>	Distribution limited to defence departments; further distribution only as approved	<input type="checkbox"/>	Other
<input checked="" type="checkbox"/>	Unlimited distribution													
<input type="checkbox"/>	Distribution limited to defence departments and defence contractors; further distribution only as approved													
<input type="checkbox"/>	Distribution limited to defence departments and Canadian defence contractors; further distribution only as approved													
<input type="checkbox"/>	Distribution limited to government departments and agencies; further distribution only as approved													
<input type="checkbox"/>	Distribution limited to defence departments; further distribution only as approved													
<input type="checkbox"/>	Other													
11. ANNOUNCEMENT AVAILABILITY (any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availability (10.) However, where further distribution (beyond the audience specified in 10) is possible, a wider announcement audience may be selected.)														
12. SPONSORING ACTIVITY (the name of the department project office or laboratory sponsoring the research and development. Include the address.) Air Command Headquarters, Westwin, Manitoba, R3J 0T0														

DSIS DCD03
HFD 09/94

UNCLASSIFIED
SECURITY CLASSIFICATION OF FORM
(Highest classification of Title, Abstract, Keywords)

UNCLASSIFIED

SECURITY CLASSIFICATION OF FORM

(Highest classification of Title, Abstract, Keywords)

13. ABSTRACT (a brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), (R), or (U). It is not necessary to include here abstracts in both official languages unless the text is bilingual).

The Defence and Civil Institute of Environmental Medicine (DCIEM) and Air Transport Group (ATG) were tasked to conduct a joint study of human factors concerning the CC-130 Hercules aircraft. The aim of the study was to establish human factors issues relevant to air accidents, and to recommend preventative measures. The study was organized around two working groups: the Crew Behaviour Assessment Group (CBAG) and the Flight Performance Assessment Group (FPAG). The CBAG developed a method of measuring the ability of the crew to coordinate their activities efficiently and manage their workload. The FPAG developed a method of measuring the accuracy and consistency of simulator flight along an aircraft flight path. Data to support the development of both methods were obtained from a simulator study of 23 ATG crews. The results defined the characteristics of high proficiency Aircraft Commanders (ACs) and those of less proficient ACs. Less proficient ACs seemed to focus primarily upon systems-related, procedural cross-checking and rechecking of information, and had more open-loop communication which supports the contention that these individuals were becoming task overloaded. The results suggest that a proportion of ATG crews are adversely overloaded by the occurrence of unexpected flight events and certain system failures. Since behaviour can be influenced by training, this study recommends a review and modification of the current CC-130 training program, including Aircrew Coordination Training (ACT).

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible, keywords should be selected from a published thesaurus, e.g. Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus identified. If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

Aircraft Human Factors
CC-130
Hercules
CFB Trenton
Training
Experience
Simulator